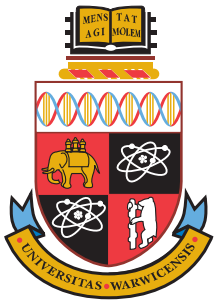


An NMR Study into the Structure and Interactions of the Transmembrane Domains of Neu and PDGFβ Receptor Tyrosine Kinases

by

Muhammad Hasan

Thesis

Submitted to the University of Warwick
in partial fulfilment for the degree of
Doctor of Philosophy

Supervisors: Dr Ann Dixon, Dr Józef Lewandowski, Prof Steven Brown and Prof Michael Allen

Molecular Organisation and Assembly in Cells Doctoral Training Centre

August 2015

THE UNIVERSITY OF WARWICK

Contents

List of Figures

List of Tables

Abbreviations

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
aa	Amino acids
ACN	Acetonitrile
ATR-FTIR	Attenuated total reflectance Fourier transform infrared spectroscopy
AU	Absorbance units
BMRB	Biological magnetic resonance databank
BPV	Bovine papilloma virus
CD	Circular Dichroism
CMC	Critical micelle concentration
CP	Cross polarisation
CSA	Chemical shift anisotropy
Da	Daltons
DARR	ADipolar assisted rotational recoupling
DEPC	1,2-dierucoyl-sn-glycero-3-phosphocholine
dH_20	distilled water
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
DPC	Dodecylphophocholine
DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
EM	Electron microscopy
ESI-MS	tElectrospray ionisation mass spectrometry
FT	Fourier transform
g	Grams
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPLC	High performance liquid chromatography
hr	hours
HSQC	Hetronuclear single quantum coherence
Hz	Hertz
IPA	Isopropyl alcohol
MAS	Magic angle spinning
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser enhancement
PDGFR	Platelet derived growth factor receptor
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
ppm	Parts per million
REDOR	Rotational echo double resonance
RF	Radio frequency
RPM	Revolutions per minute
RT	Room temperature
RTK	Receptor Tyrosine Kinase
SDS	Sodium dodecyl sulphate
ssNMR	Solid state NMR
T_1	Longitudinal relaxation time
T_2	Transversal relaxation time
TFA	Trifluoroacetic acid
TFE	2,2,2-Trifluroethanol
TM	Transmembrane
UV	Ultraviolet

Acknowledgements

First and foremost, I would like to thank my supervisors Dr Ann Dixon, Dr Józef Lewandowski, Prof Steven Brown and Prof Michael Allen for giving me the opportunity to work in an exciting area of science. Their guidance was invaluable at every stage of this research. I am deeply grateful to them for going the extra mile to ensure that I had all the support I need.

I would also like to thank everyone, in all the research groups I was a part of, for providing a great environment to work in. I would particularly like to thank the following: Mike Chow for teaching me techniques in molecular biology, Dharmesh Patel for teaching me techniques associated with peptide purification, Jonathan Lamley for helping me set up solid-state NMR experiments, Dhadchayini Jeyaharan and Leo Bowsher for support in the laboratory and amazing camaraderie throughout the past 4 years.

I would also like to thank Dr Ivan Prokes for training me on the 700 MHz spectrometer and helping me troubleshoot problems during its usage. I am also indebted to Anne Smith in the Chemical Biology Research Facility (CBRF) for ensuring all research activities in the CBRF ran smoothly and for her frequent and kind words of encouragement.

I am also indebted to Prof Alison Rodger and Dr Hugo van den Berg for giving me a place on the MOAC DTC course and helping me secure funding to enable me to stay on after my MSc. I would also like to thank everyone else at the MOAC DTC, particularly the 2010 cohort, for all the good times spent together when away for training courses and conferences, and also for encouragement and support throughout my time at Warwick University.

I would also like to thank all of my friends who provided a strong network of support throughout my PhD. I am particularly grateful to Zahra Shah and Abdurrahman Ahmad for their encouragement during thesis writing.

Special thanks goes to my parents, Naveed and Yasmin Hasan, for their unwavering love, guidance and support throughout my life. I would also like to thank my siblings and their spouses: Ayaz Hasan and Tooba Amjad, Samra Hasan and Ali Farooq, Saamia Hasan and Kasim Hasan, for their love and encouragement. Finally, l would like to thank my nephew, Hamza Ali, and nieces, Izzah Ali and Aisha Hasan, for adding so much joy to my life.

Declarations

The work in this thesis is original, and was conducted by the author, unless otherwise stated, under the supervision of Dr Ann M. Dixon (Department of Chemistry), Dr Jzef Lewandowski (Department of Chemistry), Professor Steven P. Brown (Department of Physics) and Professor Michael Allen (Department Physics).

The work was funded by the EPSRC, through the Molecular Organisation and Assembly in Cells Doctoral Training Centre.

It has not previously been presented for another degree.

Funding was provided by an EPSRC studentship.

All sources of information have been acknowledged by means of reference.

Abstract

Receptor tyrosine kinases (RTKs) form a medically important class of membrane proteins whose malfunction as signalling entities has been reported as the cause of a large number of diseases. The structure and dynamics of the transmembrane domain (TM) of these receptors is central to their role in cellular signalling. However, this structure–function relationship is poorly understood. This is largely due to the problems associated with the production of these proteins for analyses *in vitro* and the fact that they are hydrophobic and require membrane mimetics to solubilise them, which in turn often means loss of structural integrity.

This study aimed at using solution- and solid- state NMR to probe the structure and interactions of two biologically important RTK-TMs for which no high resolution structures exist. For the oncogenic Neu receptor tyrosine kinase transmembrane domain, it was sought to understand which amino acids are involved in stabilising the helical dimer. Synthetic peptides were reconstituted into liposomes and then analysed using solid-state NMR. The platelet-derived growth factor β receptor transmembrane domain was another RTK-TM whose structure and interactions with E5 protein were studied using solution-state NMR. A detailed protocol for the expression, purification and reconstitution of PDGFβR-TM into membrane mimetics was devised in order to perform the stated NMR analyses.

Chapter 1

Introduction

1.1 Membrane Proteins

Any living cell houses a host of constituents which are compartmentalised using a barrier known as the plasma membrane. Furthermore, organelles within a cell partition certain cellular machinery using membranes. These biological membranes are made up of a variety of phospholipids (mostly replaced by glycolipids in algae and plant chloroplasts) which self-associate to form a bilayer. Embedded within these bilayers, amongst other biological molecules (e.g cholesterol), are a class of proteins known as membrane proteins. These proteins are involved in almost all aspects of cellular function, from relaying information from the external environment to the cell, to trafficking material in and out of the cell. As such, their malfunction usually has serious consequences and they have been implicated in a broad range of diseases such as cancer, inflammatory and neurological diseases.

Genomic studies over the last couple of decades have revealed that membrane proteins constitute as much as a third of eukaryotic proteins (Baker, 2010) and investigating their structure and function has never been more important. Even though membrane protein structure determination has moved forward leaps and bounds within the last decade due to novel techniques in X-ray crystallography, NMR spectroscopy, and more recently, cryo-electron microscpy, they still represent less than 3% of the structural entries in the Protein Data Bank (PDB) (Berman et al., 2000). This is in stark contrast with their pharmacological importance given that around 60% of the current approved drug targets are membrane proteins (Yıldırım et al.,

2007).

Membrane proteins have traditionally presented themselves as very challenging systems to study for structural biologists. This is primarily because they natively reside within the hydrophobic environment of the lipid bilayer and their analysis requires a membrane-mimicking environment. This is usually achieved by studying the protein *in vitro* in the presence of detergents and/or lipids, which exponentially increases the obstacles that need to be overcome for successful high-resolution structural analysis (Warschawski et al., 2011). Furthermore, membrane proteins are present natively in the membrane only at very low concentrations while very often milligram quantities are required for structural analyses. This problem is usually overcome by overexpressing the protein of interest recombinantly in heterologous host organisms. Although *Escherichia coli* has been extensively employed for this purpose, other bacterial hosts (e.g. *Lactococcus lactis*), yeast expression systems (e.g. *Picha pastoris* and *Saccharomyces cerevisiae*), insect and mammalian systems are also routinely used. Extensive and systematic screening of expression systems often needs to be performed to obtain a membrane protein in sufficient quantities while maintaining structural integrity.

1.2 Receptor Tyrosine Kinases (RTKs)

RTKs are cell-surface phosphotransferase enzymes that specifically phosphorylate the tyrosine amino acid using ATP and create ADP and a phosphotyrosine in the protein. This allows protein substrates to bind to docking sites in the kinase domain and hence allow RTKs to mediate transmembrane transduction of extracellular signals to the cytoplasm for the downstream control of various cellular processes including cell-proliferation, differentiation, metabolism and motility (Blume-Jensen and Hunter, 2001; Ullrich and Schlessinger, 1990).

The overall topology of RTKs is highly conserved with a ligand-binding region in the extracellular domain, a single transmembrane helix, and a cytoplasmic region which includes a regulatory juxtamembrane domain and a catalytic kinase domain (Lem-

mon and Schlessinger, 2010) (Figure 1.1). The ligands for RTKs include diffusible peptides or small proteins such as growth factors, cytokines, hormones, neurotrophic factors, and other extracellular signalling molecules (Li and Hristova, 2010).

The traditional model for the mechanism of RTK activation involves ligand binding to the extracellular domain which induces dimerisation with an adjacent receptor, which in turn causes the kinase domain to become active. The activated receptor then autophosphorylates on multiple intracellular tyrosine sites. The phosphorylated tyrosines can then bind to specific substrates. These may be enzymes or molecules that can bind to enzymes elsewhere. The products generated from this process may yet bind elsewhere and thus continuing the signalling cascade (Lemmon and Schlessinger, 2010). This is illustrated in Figure 1.1.

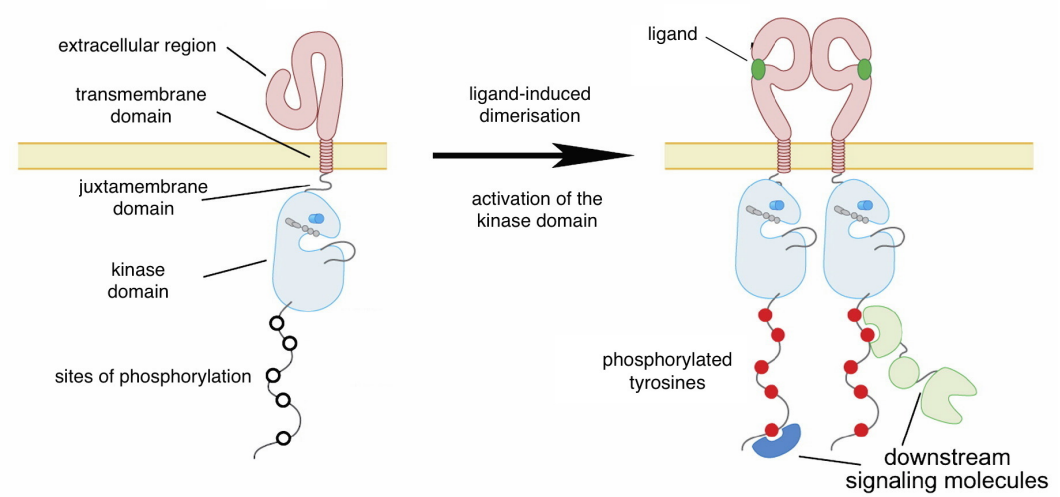


Figure 1.1: Schematic for ligand-induced activation of RTKs. Adapted from (Zhang et al., 2006).

An alternative model for receptor activation which has been garnering increasing evidence is that of rotation-coupled activation. Here, the receptors exist as preformed dimers and ligand binding induces a rotation of the transmembrane domain such that the kinase domain is conferred a conformational rearrangement to an active, asymmetric state. Rotation of the transmembrane helices and variations against the tilt angle are also more energetically feasible than lateral association of the monomer receptors in the lipid bilayer, as proposed by the ligand-induced activation model. Additionally, the rotation-coupled activation model can also explain bidirectional

signal transduction, where cytoplasmic factor(s) may induce transmembrane rotations in reverse direction to those of 'outside-in' signalling (Maruyama, 2015).

Mutations in RTKs and other biological processes which result in aberrant intracellular signalling have been implicated in a wide variety of diseases such as cancers, diabetes, inflammation, severe bone disorders and angiogenesis. Their increasing medical significance has provided an impetus for the development of drugs which block or attenuate RTK activity (Lemmon and Schlessinger, 2010). As such, investigations of the structure–function relationship of RTKs has become more important than ever. This is particularly true for the transmembrane domains of these receptors, which have been particularly challenging for high-resolution structure determination methods, with only two high-resolution structures of transmembrane domain dimers available (Bocharov et al., 2008b,a). This study will investigate the structural properties of two RTKs: Neu and PDGFβR.

1.2.1 Neu RTK

The ErbB family of RTKs have been shown to play vital roles in cardiac, neurological and endocrine systems (Bublil and Yarden, 2007). The activation mechanism of this family is of considerable medical interest because mutations and deletions that result in aberrant signalling have been identified in a number of human tumours (Lemmon and Schlessinger, 2010; Holbro et al., 2003). The constitutive activity of mutant forms of ErbB2/Her2 receptor in particular have been implicated in a very aggressive form of breast cancer which forms about 30% of all human breast cancers, and in many other cancers such as ovarian, stomach, bladder, salivary, and lung carcinomas (Tan and Yu, 2007). The involvement of ErbB2/Her2 in various cellular processes is shown in Figure 1.2.

The rat homologue of this receptor, Neu, has been shown to undergo full oncogenic transformation following a $V_{664}E$ mutation in the transmembrane domain. Studies conducted on the Neu RTK have provided evidence that specific interactions between transmembrane helices can mediate dimerisation and receptor activation (Hynes and

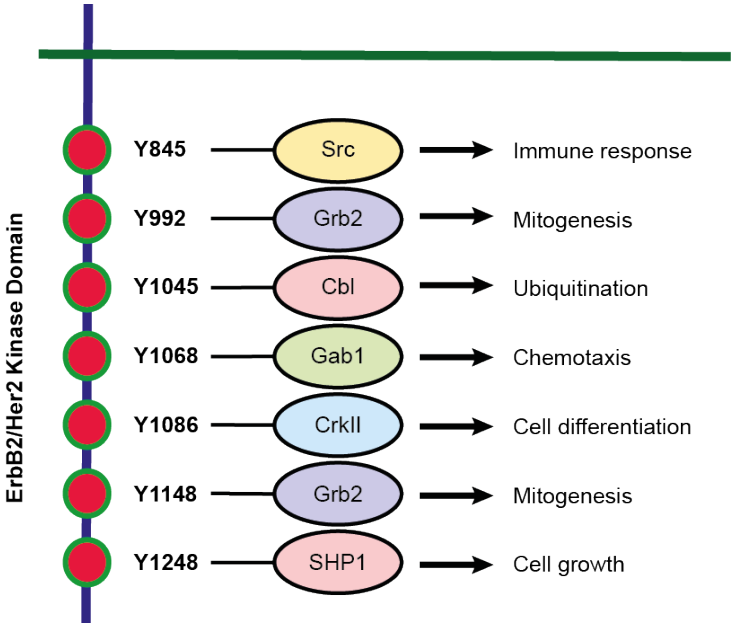


Figure 1.2: The involvement of various Neu kinase domain tyrosines in cell signalling processes. Examples of substrates that bind to specific tyrosines are given in ovals.

Stern, 1994; Bell et al., 2000; Smith et al., 2002a) and hence it provides a valuable paradigm for investigating the mechanistic details of RTK transmembrane domains in signal transduction.

1.2.2 Platelet derived growth factor receptor-β (PDGFβ) RTK

Platelet derived growth factors (PDGFs) are a family of four cystine-knot-type growth factors (PDGF-A, -B, -C and -D) which have important roles in regulating cell proliferation, differentiation, growth and development (Williams, 1989). There are two forms of receptors for PDGFs: PDGFαR and PDGFβR. Both of them belong to the class III subfamily of RTKs, and others in the family are the stem cell growth factor receptor (Kit), macrophage colony-stimulating factor (M-CSF) receptor, and FMS-like tyrosine kinase 3 (Schlessinger, 2000). All of these receptors have very important physiological roles; PDGFβR in particular is an essential regulator of early hematopoiesis and blood vessel formation (Andrae et al., 2008) through chemotactic and mitogenic signalling (see Figure 1.3). Malfunction of the receptor has been linked to diseases such as inflammation, pulmonary fibrosis and restenosis,

atherosclerosis and many types of cancer.

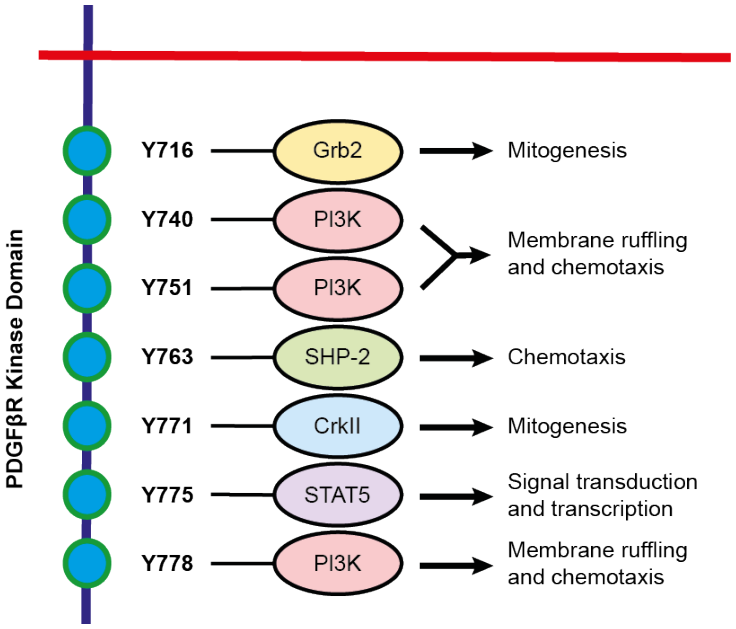


Figure 1.3: The involvement of various PDGFβR kinase domain tyrosines in cell signalling processes. Examples of substrates that bind to specific tyrosines are given in ovals.

Structural studies have shown that PDGFβR transmembrane (TM) domains have the propensity to dimerise on their own (Oates et al., 2010; Muhle-Goll et al., 2012). The smallest known oncoprotein, E5 from bovine papillomavirus type-1, is known to induce ligand-independent activation of the receptor. This interaction will be discussed in detail in section 4.1 and involves the E5 homodimer forming a complex with PDGFβR-TM slowing down receptor internalisation. This in turn generates a sustained proliferative signal resulting in tumorigenic transformation (Nilson and DiMaio, 1993).

1.3 Membrane Mimetics

Biological membranes are essential in separating the constituents of a cell or organelle from their external environment. However, the transfer of information, nutrients and other biological molecules through these membranes is key in order to retain proper biological function. As such, membranes are a complex heterogeneous barrier

made up of not only lipids but also a variety of other biomolecules such as proteins, carbohydrates and cholesterol.

Membrane proteins need to be extracted from their natural environment and reconstituted in artificial milieus for the structural analysis by X-ray crystallography or NMR. The choice of membrane mimetics should, in general, satisfy two criteria: first, it should replicate the native environment of the protein as closely as possible as protein–lipid interactions may play a key role is structure and function; second, it should be compatible with downstream applications and also allow data acquisition of the highest quality. The search for membrane mimetics that conserve the native structure and function remains a prime challenge for structural biologists. This section will discuss some of the most commonly used membrane mimetics with a particular focus on their application to NMR spectroscopy.

1.3.1 Detergent micelles

Detergents are the most commonly used membrane mimetics to characterise membrane proteins in an aqueous medium (Kang and Li, 2011). These are amphiphilic molecules with a hydrophilic 'head' and a hydrophobic 'tail'. Above a certain concentration, known as the critical micellar concentration (CMC), detergent monomers self-associate to form micelles where the hydrophilic head is solvent-exposed and the hydrophobic tails are in contact with each other (Figure 1.4).

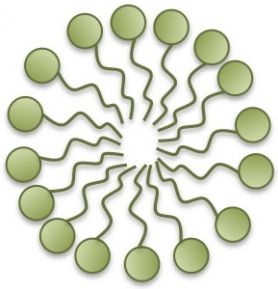


Figure 1.4: Representation of an assembly of detergent molecules to form a micelle. The polar head groups face outwards towards the solvent whereas the acyl chains cluster together due to their hydrophobicity.

Detergents are an invaluable tool in a wide array of biochemical and biophysical

techniques, including X-ray crystallography and solution-state NMR spectroscopy. Their significance in the latter is underscored by the fact that a large majority of membrane protein structures obtained to date have been in detergents. This is primarily because detergent micelles are much smaller (∼30–50Å) and hence tumble faster on the NMR timescale which allows the acquisition of high-resolution NMR spectra (Warschawski et al., 2011). The challenge when screening detergents for NMR analyses is to find a detergent which confers to the protein a high degree of conformational stability, does not impose curvature stress which would greatly alter protein structural integrity and gives the protein–detergent complex the ability to withstand lengthy periods of data collection, often at elevated temperatures.

Despite the advantages that detergent micelles bring in solubilising membrane proteins for structural studies, a preferential system would be where the environment more closely resembles the lipid bilayer that the membrane protein is embedded in *in vivo*. Although recreating the complexity of native membranes is not possible *in vitro*, the addition of lipids to the micelle (forming what is known as a 'mixed micelle') has proven very useful in analysing membrane proteins using solution-state NMR. For example, a recent study investigating Yop1p protein found that helical propensity was increased at the edges of the TM domain and also more homogeneous signals were observed in general (Brady et al., 2015). Another example is when studying mammalian retinal opsin protein it was found that it is generally very unstable in detergents alone; however, stable opsin was obtained when purified in a mixture of lipids and detergents (Reeves et al., 1999).

1.3.2 Liposomes

Lipid vesicles, or liposomes, are formed when lipids organise themselves as bilayers in an aqueous environment (Figure 1.5). As they are made up of phospholipid bilayers, they very closely resemble native conditions. Although lipids spontaneously tend to form non-homogeneous multilamellar vesicles, a homogeneous size distribution can often be obtained through extrusion. Additionally, properties of the membrane such as fluidity, length and head group charge can be manipulated to further resemble

native conditions. Membrane proteins are usually inserted into the liposomes using a detergent-mediated pathway as detergents are convenient for the initial solubilisation, isolation and purification of the protein. There are a variety of methods to subsequently insert the protein in to liposomes but they generally involve dialysis or incubation with detergent-adsorbing beads (Seddon et al., 2004).

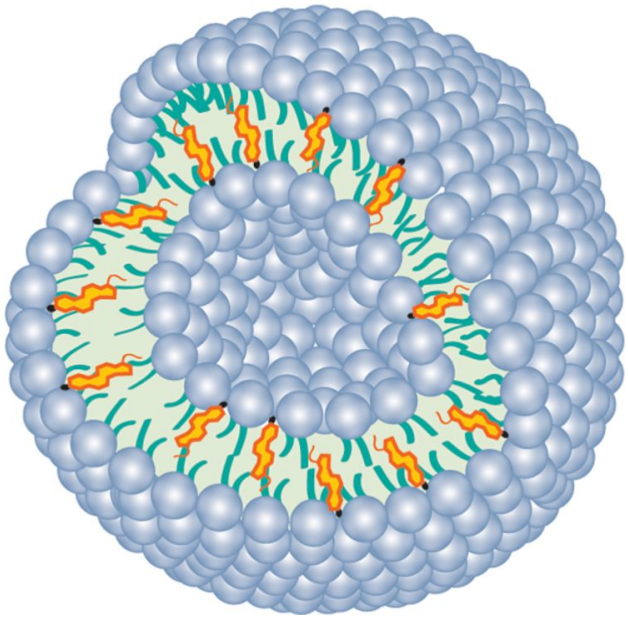


Figure 1.5: A diagrammatic illustration of a liposome. Lipid molecules arrange themselves in a bilayer with the headgroups facing towards the solvent. A heterogeneous mixture of lipids and other biomolecules such as cholesterol can be used to make liposomes which mimic more native-like conditions. Adapted from (Monteiro et al., 2014).

Liposomes are excellent mimetics for membrane bilayers but their large size precludes them from analysis through solution-state NMR. However, it is this same property which makes them amenable for solid-state NMR analysis. Static NMR spectra of liposomes generally suffer from very poor resolution but crucial information involving protein–lipid interaction can be obtained through ^{31}P NMR and/or ^{2}H NMR when using deuterated lipids (Ali et al., 2001).

Significant improvements in resolution of solid-state NMR spectra can be achieved by spinning the sample at the magic angle. Magic angle spinning solid-state NMR (MAS ssNMR) on liposomes has frequently been the technique of choice for a number of structural investigations over the last decade and has provided unprecedented

insights into membrane protein structure, for example, the Influenza M2 channel in DLPC or DMPC liposomes (Cady et al., 2010) and protegrin in POPC liposomes (Mani et al., 2006).

As mentioned earlier, although liposomes are excellent membrane mimetics, they do not completely capture the complexity of a native biological membrane. The effect of lipid composition on membrane protein structure has been gaining increasing importance recently (Lee, 2004; Dowhan and Bogdanov, 2011). The commercial availability of a heterogeneous mixture of lipid extracts from *E. coli* and yeast etc. has greatly aided several membrane protein studies (Rice et al., 2014; Brady et al., 2015). Furthermore, recent advances in solid-state NMR methodology, such as the development of Dynamic Nuclear Polarisation (DNP), have been exploited to demonstrate the feasibility of studying overexpressed membrane proteins in native *E. coli* membranes (Fu et al., 2011; Renault et al., 2012; Ward et al., 2015).

1.3.3 Bicelles

Bicelles (**bi**layered mi**celles**) are formed when, at the correct concentration and temperature, certain detergents such as DHPC or CHAPS are mixed with lipids such that they assemble into a bilayered discoidal structure (Sanders and Prosser, 1998) (Figure 1.6). Bicelles have prove to be a valuable tool for studying integral membrane proteins using solution-state NMR. As they represent an intermediate morphology between micelles and liposomes, they maintain the attractive feature of providing a bilayer environment while being monodisperse and small enough to be suitable for solution-state NMR analysis. Bicelles have been instrumental in performing structural studies on membrane proteins (Lau et al., 2008; Bocharov et al., 2007, 2008a).

Depending on downstream use, bicelles can be prepared in two different sizes. Small isotropic bicelles which tumble rapidly in solution are suitable for solution-state NMR. Larger bicelles can also be magnetically aligned with the magnetic field and can be used for studying orientation and crossing angles of membrane proteins em-

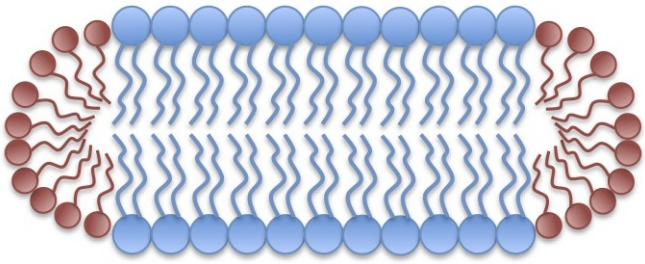


Figure 1.6: Representation of a bicelle. Short chain detergent molecules (red) form a rim around the long chain lipid molecules (blue).

bedded within bicelles (Sanders and Prosser, 1998). Not only can aligned bicelles be used in the solution state, but solid-state NMR can also be used to investigate protein structure in aligned bicelles. An example is when a combination of solution- and solid-state NMR was used to characterise local and global dynamics of the chemokine receptor CXCR1 (Park et al., 2006).

Bicelles have also proven to be useful in other arenas of structural biology. The helical tilt of phospholamban, a regulatory single-pass transmembrane protein, and its segmental mobility were investigated by electron paramagnetic resonance (EPR) spectroscopy in oriented bicelles (Ghimire et al., 2012). X-ray crystallography has also employed bicelles to successfully solve several membrane protein structures (Faham and Bowie, 2002; Rasmussen et al., 2007; Payandeh et al., 2011) and as such the arrival of bicelles has had far-reaching implications in structural biology.

1.3.4 Nanodiscs and SMALPS

Nanometric lipid bilayers, or nanodiscs, form a planar phospholipid bilayer like bicelles but instead of detergents, the lipids are encompassed by a dimer of a lipoprotein known as membrane scaffold protein (MSP) (Figure 1.7) . During the last 10 years, a large number of solution-state NMR studies have been performed on membrane proteins embedded in nanodiscs and a number of high-resolution spectra have been obtained (Tzitzilonis et al., 2013; Shenkarev et al., 2013; Yokogawa et al., 2012). The first membrane protein structure solved using nanodiscs was that of OmpX which highlighted the potential of these mimetics (Hagn et al., 2013). One

major advantage of using nanodiscs is that various MSPs can be engineered to define the size of the nanodisc. This allows membrane proteins to be reconstituted in a more controlled manner and factors such as the oligomeric state can be manipulated (Ritchie et al., 2009). This benefit was exploited in the study of EGF receptor tyrosine kinase where they showed RTK kinase activity was significantly stabilised in nanodiscs (Mi et al., 2008).

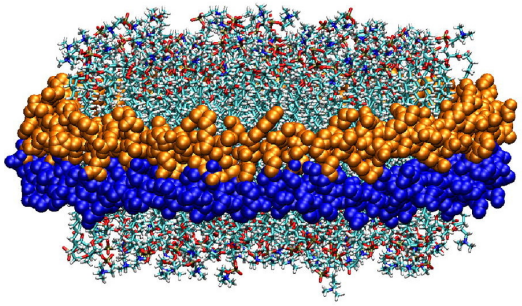


Figure 1.7: Illustration of a nanodisc structure. The two membrane scaffold proteins (coloured orange and blue) for a ring-like cassette which encompasses the lipid bilayer. Adapted from (Bayburt and Sligar, 2010).

Nanodiscs have also been used in solid-state NMR studies in two different ways. One approach is to use large nanodiscs, also known as macrodiscs, which can be precipitated with polyethylene glycol (PEG) or lyophilisation and then packed into an MAS rotor (Li et al., 2006; Mörs et al., 2013). Since nanodics are monodisperse, membrane proteins obtained in this fashion are expected to be more homogeneous and hence provide narrower NMR linewidths. More recently, nanodsics have been packed into MAS rotors by through centrifugal sedimentation (Ding et al., 2015). This avoids issues with precipitation and lyophilisation which may negatively affect protein structure and function.

There are several synthetic polymers which have shown significant promise in membrane protein structural analyses. One such polymer is styrene maleic acid lipid particle (SMALP). SMALPs are made up of alternating styrene (hydrophobic) and maleic acid (hydrophilic) groups; this property allows them to insert into a biological membrane and extract discs of the lipid bilayer. Unlike amphipols, bicelles and nanodiscs, SMALPs can bypass detergent extraction and subsequent reconstitution making it a very powerful method for *in vivo* membrane protein analysis.

Although several studies have been used to analyse membrane proteins using this method (Knowles et al., 2009; Jamshad et al., 2011; Postis et al., 2015), a thorough characterisation is still lacking and some studies have shown these SMALPs can significantly alter the physical properties of the membrane (Orwick et al., 2012).

1.3.5 Amphipols

Amphipols are short amphipathic polymers that can solubilise integral membrane proteins (Tribet et al., 1996). For certain membrane proteins, amphipols are favoured over conventional detergent micelles because of their ability to increase stability of the embedded membrane protein while exhibiting good refolding properties (Elter et al., 2014). Amphipols possess great potential for solution-state NMR studies owing to the fact that upon solubilisation of membrane proteins, very few polymer molecules remain disassociated with the amphipol–protein complex overcoming problems associated with viscosity increase at high concentrations (Seddon et al., 2004). Having said that, problems to do with polymer aggregation and deleterious effects of the amphipol scaffold on protein function have been reported (Picard et al., 2006). Like bicelles, amphipols have also been used to determine high resolution structures through X-ray crystallography (Polovinkin et al., 2014).

1.4 NMR as a tool for the elucidation of membrane protein structures and interactions

The advent of novel sample preparation techniques coupled with developments in NMR methodologies, both in solution and solid state, have made an increasing number of membrane proteins amenable to comprehensive structural studies. NMR structures of integral membrane proteins constitute 10–15% of overall structures (Nietlispach and Gautier, 2011). This proportion has the potential to increase as limiting aspects continue to be overcome with the employment of techniques such as the use of lanthanide-induced pseudo contact shifts in solution-state NMR (Crick

et al., 2015), and the use of detergent microcrystals in solid-state NMR (Shahid et al., 2012a).

NMR spectroscopy is uniquely placed in the field of structural biology in that it is able to report on protein dynamics at atomic resolution over a wide range of timescales. The importance of studying protein dynamics is underscored by the fact that a host of biological processes such as enzymatic activities, protein folding, and cellular signalling rely on time-dependent fluctuations in conformation. A recent study, which illustrates the power of NMR spectroscopy in studying these dynamics, reports a direct observation of hierarchy of protein and solvent motions in the protein conformational energy landscape (Lewandowski et al., 2015). Solution- and solid-state NMR methods have frequently been used to study protein dynamics in membrane mimetics as well. For example, ^{15}N relaxation measurements have been used to study relative motions between helices of polytopic membrane proteins (Hwang et al., 2004; Villinger et al., 2010). In this vein, NMR spectroscopy is an excellent method which can be used in conjunction with other high-resolution methods which can only provide a static picture of membrane protein structure.

Membrane mimetics used to determine the high-resolution structure of proteins embedded in a bilayer have been highlighted in the previous section. However, most of the full length structures that have been obtained are of integral membrane proteins and structures of single-pass membrane proteins are rare as conditions required for a stable structural fold are extremely difficult to obtain, as has been pointed out previously (Yamamoto et al., 2015). This study focuses on the transmembrane (TM) domains of RTKs, and while no high-resolution structure of a full length RTK exists, only two high resolution structures have been obtained of RTK-TM dimers and both were obtained through the use of solution-state NMR spectroscopy (Bocharov et al., 2008b,a). New sample preparation methods specifically designed to study these TM domains (Bugge et al., 2015), as well as novel NMR methodologies can be combined to give an unprecedented insight into the structure and function of these RTKs.

1.5 NMR Theory

1.5.1 Spin

Spin, or spin angular momentum, is a quantum mechanical property which is intrinsic to all nuclei that do not contain an even number of protons and neutrons. Spin is characterised by a spin quantum number (I) which can either be an integer or half integer. If the spins in an atom are paired (e.g. in ^{12}C and ^{16}O), then they will cancel each other out ($I = 0$) rendering the nucleus NMR-inactive.

All other spins produce a magnetic dipole moment (μ) related to the spin quantum number through the equation:

$$\mu = \gamma\hbar\sqrt{I(I+1)}$$
$$\text{where } \hbar = \frac{\text{Planck's constant}}{2\pi}.$$

The constant of proportionality (γ) is known as the gyromagnetic ratio and is unique to each nuclear isotope. When the nuclear spins are placed in an external magnetic field (B_0), they will reorient themselves either with or against the magnetic field. Protein NMR spectroscopy is primarily concerned with three NMR active nuclei: ^{1}H, ^{13}C and ^{15}N. All of these have a spin quantum number of $\frac{1}{2}$, i.e. $I = \frac{1}{2}$, and as such split into two energy levels when placed in a magnetic field, one parallel and and one anti-parallel (Figure 1.8). The distribution of these nuclear spin energy levels is governed by the Boltzmann distribution from which it is possible to deduce that at thermal equilibrium, the two energy levels are nearly equally populated. Nevertheless, the very small excess in the lower energy state which is of consequential importance as it is this net magnetisation (M_0) which can be measured, albeit the low difference in population levels renders NMR a method with inherently weak sensitivity.

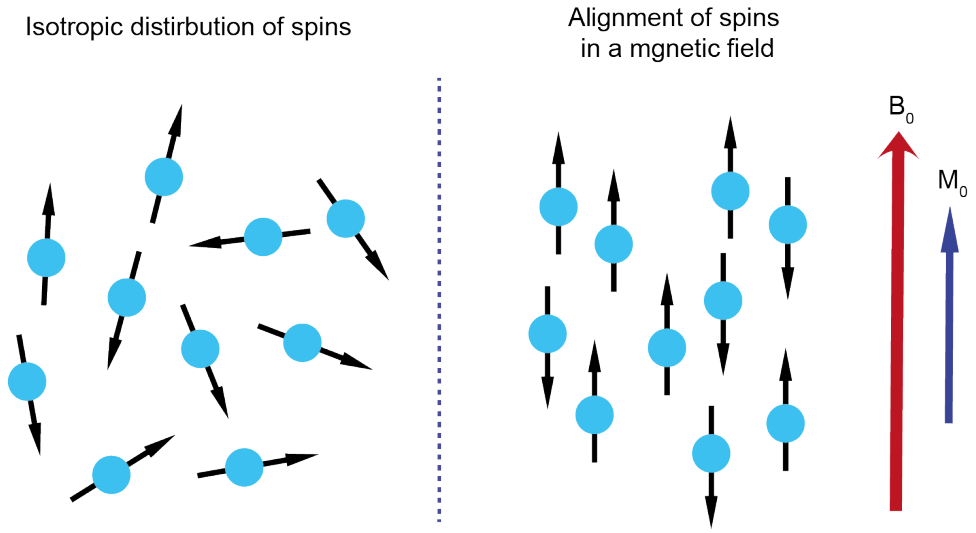


Figure 1.8: Alignment of spins in a magnetic field. When spins are placed in a magnetic field (B_0), they are no longer isotropically distributed and align themselves either with or against the field. There is a small excess of parallel spins resulting in a net magnetisation (M_0).

1.5.2 Larmor precession

When a nucleus with spin angular momentum is placed in a magnetic field (B_0) applied along the z-axis, the magnetic moment will precess along this axis as shown in Figure 1.9. The rate of precession is known as the Larmor frequency (ω_0) and its dependence on the gyromagnetic ratio (γ) and the magnetic field (B_0) is defined by the following equation:

$$\omega = -\gamma B_0$$

where negative sign denotes the direction of motion of the magnetic moment about the static field.

In an actual sample being analysed, there will be an extremely large number of nuclear spins all precessing about the z-axis. The net bulk magnetisation (M_0) arises from an ensemble of nuclei of the same kind precessing with the same angular frequency (Larmor Frequency) in a magnetic field.

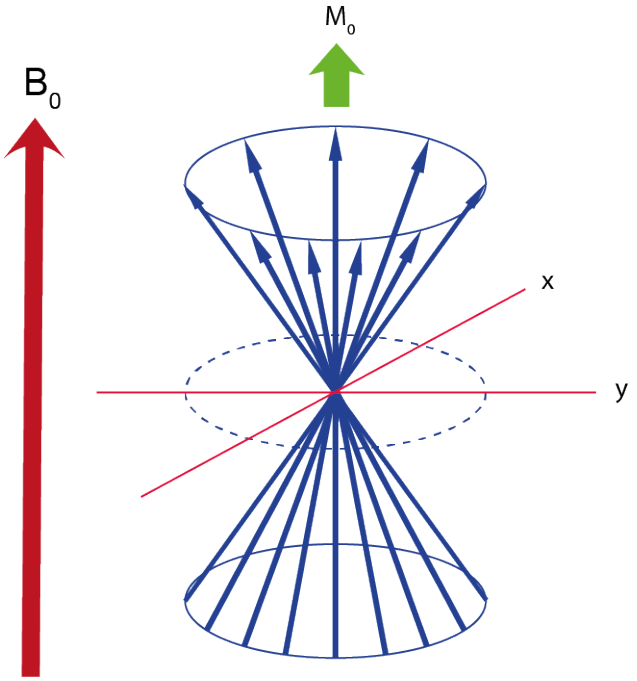


Figure 1.9: Spin precession about a magnetic field (B_0) along the z-axis. Net magnetisation (M_0) will be parallel to the magnetic field as a result of excess population of spins in the lower energy level.

1.5.3 Radiofrequency (RF) pulse

In order to detect nuclear magnetic resonance in a spectrometer, the magnetic moment must be perturbed away from the z-axis. This is done by applying electromagnetic radiation, i.e. RF pulse, perpendicular to the magnetic field at the Larmor frequency. When applied for a sufficient amount of time, this rotates the magnetisation from the z-axis to the xy-plane as shown in Figure 1.10. This pulse is called a $\frac{\pi}{2}$-excitation pulse. The precession of the magnetisation in the xy-plane can now be detected by a receiver coil through electromagnetic induction.

1.5.4 Free induction decay (FID)

Once the RF pulse is switched off, the bulk magnetisation then returns to thermal equilibrium parallel to the z-axis (Figure 1.11). This phenomenon depends on the longitudinal relaxation (T_1) and the transverse relaxation (T_2). T_1 describes the return to thermal equilibrium as a result of the interaction of spins with the sur-

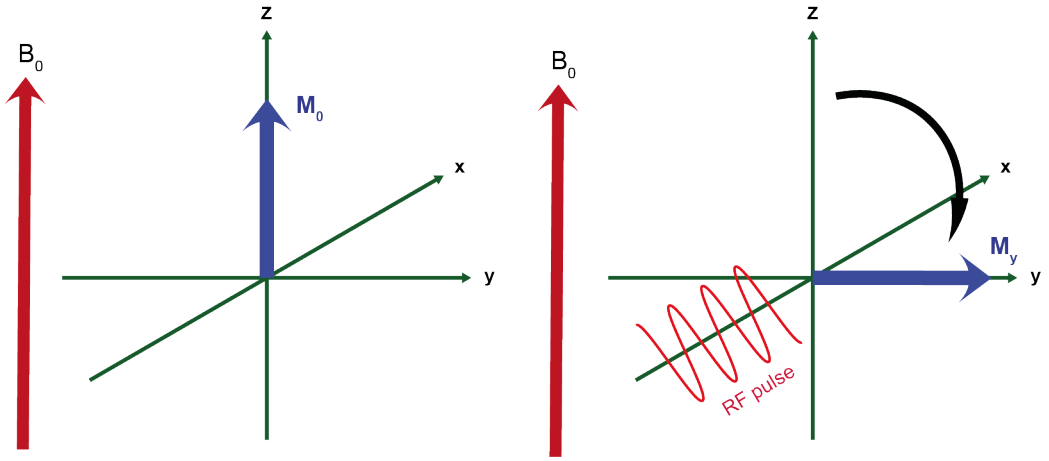


Figure 1.10: Rotation of bulk magnetisation by RF pulses.When RF pulses are applied along the x-axis, the bulk magnetisation (M_0) rotates and is transferred to the xy-plane. The precession in this plane induces an electric current which is then detected.

rounding lattice while T_2 is a result of the interactions between spins. As a result of these two relaxation rates, the detected signal decays over time (i.e. the amplitude reduces) (Figure 1.11). This is known as free induction decay and is a measure of decay rate as a function of time.

The FID contains contributions from all the different target nuclei in the sample. In order to obtain an NMR spectrum, Fourier transformation needs to be performed on the time dependent FID to yield a frequency domain spectrum.

1.5.5 Chemical Shift

The magnetic field experienced by each nucleus in a particular sample is different and results in each nucleus to precess at a distinct frequency resulting in separation of signals in an NMR spectrum. This is because each nucleus in the sample produces a local magnetic field which opposes the external field. This local magnetic field is determined by the local electronic environment. It is due to this that nuclei of the same type in different chemical environments experience slightly different local magnetic fields and therefore have different frequencies of precession. The chemical shift can be defined as the difference between a frequency of interest and a reference frequency of a chosen standard. Hence, nuclei in a particular sample can

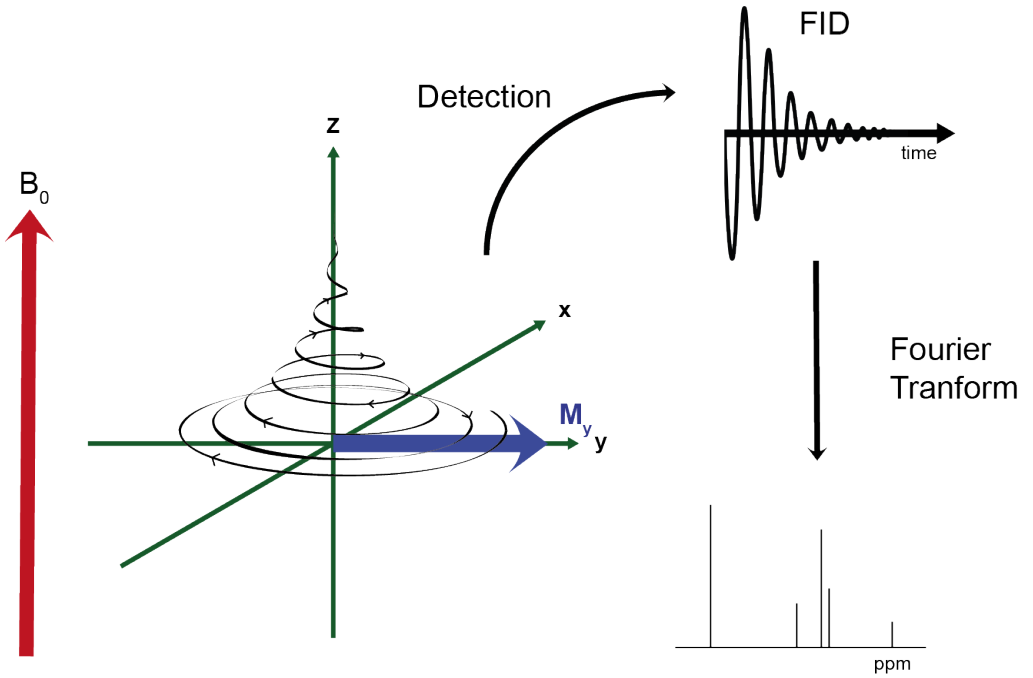


Figure 1.11: Free induction decay.

be differentiated from each other by virtue of their chemical shift.

In order to remove the reliance of the chemical shifts on the magnetic field strength, the chemical shifts are converted from Hz to ppm using the following equation:

$$\delta_{\text{ppm}} = \frac{\omega_0 - \omega_{\text{ref}}}{\omega_{\text{ref}}} \times 10^6$$

where ω_{ref} is the Larmor frequency of the chosen standard e.g. tetramethylsilane (TMS) or 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). This results in a scaled spectrum that is independent of the external magnetic field (B_0) and can be used for comparative purposes.

1.5.6 Two dimensional (2D) NMR spectroscopy

Although nuclei in different chemical environments will provide distinct chemical shifts, sometimes in the case of large and/or complex molecules such as proteins there may be a crowding or overlapping of signals which can make the resulting spectrum complicated or even impossible to interpret. For this reason, multidimensional experiments are performed to successfully assign chemical shifts.

Several homonuclear and heteronuclear 2D experiments exist in solution- and solid-state NMR. However, the general scheme involves exciting all nuclei simultaneously creating a magnetisation in the xy-plane. This magnetisation is then allowed to evolve for a given amount of time (t_1). Next, during a mixing time (t_{mix}), magnetisation is allowed to transfer between nuclei using a combination of pulses. Finally, during the detection period, FID signal from the sample is recorded. A set of different FIDs is recorded by varying (t_1). A 2D NMR spectrum can then be obtained through Fourier transformation of the data. A simplified schematic of this is shown in Figure 1.12.

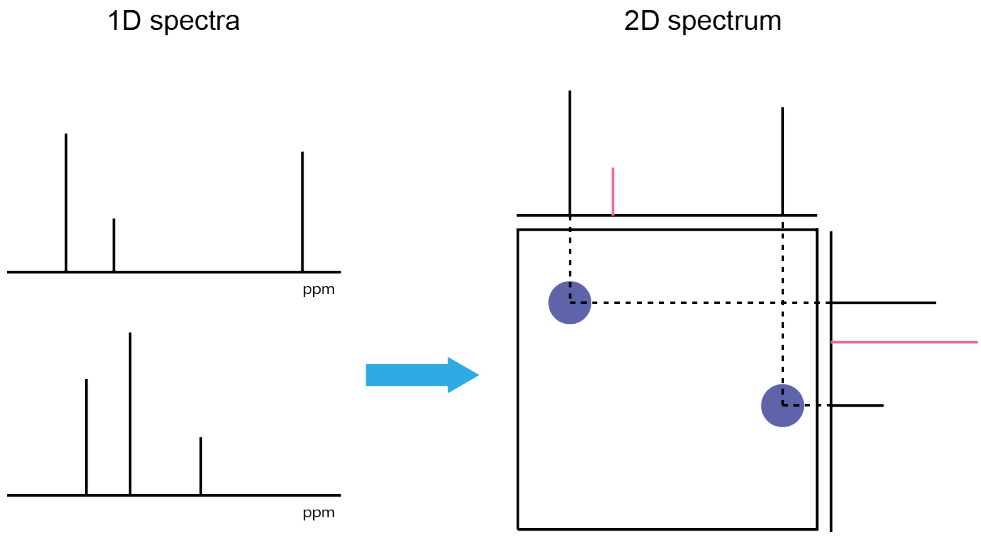


Figure 1.12: A simplified schematic of how a 2D spectrum is obtained from a series of 1D data. When there is a correlation between certain peaks in the 1D spectra, arising from magnetisation transfer, they will show up on the 2D spectrum (blue peaks in 2D spectrum). When there is no such correlation, there will not be any corresponding peak in the 2D spectrum (pink peaks in 1D spectra).

One of the most important and useful 2D experiments in biomolecular solution-state NMR is the HSQC experiment, which will be further explained in section 6.2. Similarly, a commonly used 2D experiment in protein solid-state NMR, and one that this study has frequently employed, is the DARR experiment explained further in section 1.6.4.

1.5.7 Three dimensional (3D) NMR for protein backbone assignment

Two dimensional experiments can also be extended to a third dimension. In protein NMR, backbone assignment is usually accomplished using a set of triple resonance experiments. In order to do so, a ^{1}H–^{15}N HSQC correlation spectrum is extended into a third ^{13}C dimension as illustrated in Figure 1.13. Signal overlap in the HSQC can hence be resolved and ^{13}C shifts can help sequentially link the backbone, in addition to helping identify amino acid type. For instance, serine, threonine and alanine have characteristic $^{13}C_{\beta}$ chemical shifts which makes them easily identifiable. On the other hand, glycine has no $^{13}C_{\beta}$ chemical shift.

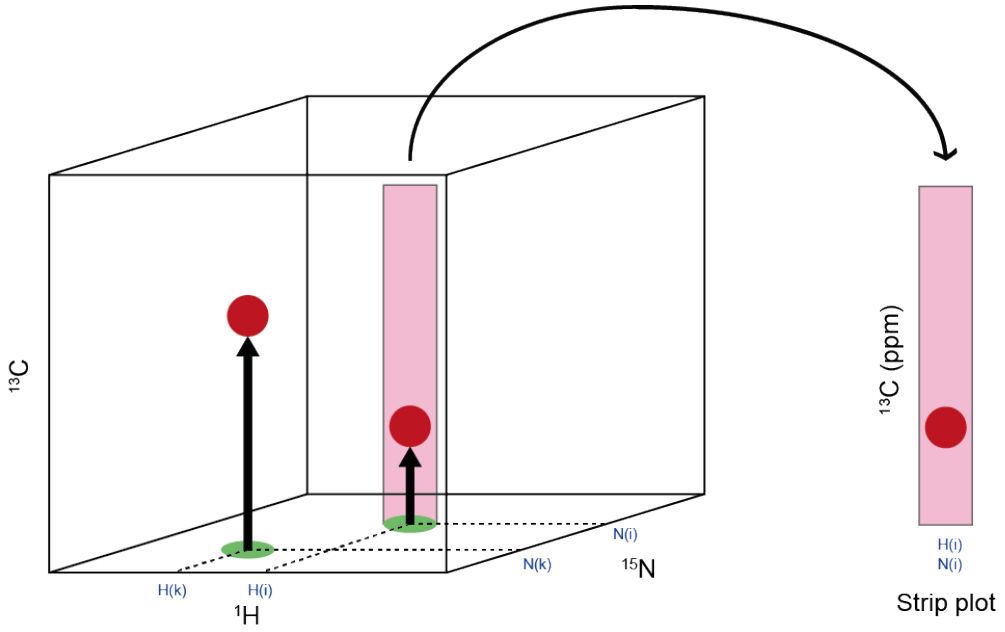


Figure 1.13: A simplified schematic of how a 2D spectrum is extended to a 3D spectrum. In this example, an HSQC experiment (green) is extended vertically using carbon as an additional dimension. The resulting 3D data set can be analysed using 'strip plots' where an entire carbon dimension is shown at a single ^{15}N and ^{1}H chemical shift. This enables us to look at carbon chemical at shifts at known values of ^{15}N and ^{1}H chemical shifts, H(i) and N(i) in the given illustration.

Most triple resonance experiments involved in backbone assignment are named according to the atoms involved in the transfer, and the sequence in which they are involved. For example, the CBCACONH experiment (Grzesiek and Bax, 1992), correlates the backbone amide hydrogen and nitrogen atoms of a given residue (i) with

the Cα and Cβ of the preceding residue $(i-1)$, as illustrated diagrammatically in Figure 1.14(a). The magnetisation is transferred via the carboxyl group but this is not observed in the spectrum as the chemical shift is not evolved on the COs. In the resulting data set, two carbon peaks are observed at the same ^{1}H and ^{15}N shifts as each peak in the HSQC.

Triple resonance experiments for backbone assignment are always performed in pairs for sequential linking of amino acids. This is necessary in order to unambiguously identify (i) and $(i-1)$ residues. The complementary pair for the CBCACONH experiment is the CBCANH experiment (Wittekind and Mueller, 1993). As shown in Figure 1.14(a), CBCANH correlates the amide hydrogen and nitrogen in the backbone of given residue (i) with the Cα and Cβ of that residue as well as the Cα and Cβ of the preceding residue. This results in four carbon peaks for given peak at a particular ^{15}N and ^{1}H chemical shifts in the HSQC. However, due to the fact that signals arising from Cβ are opposite in phase, they will appear as negative peaks. This property allows peaks from each of the residues, (i) and $(i-1)$, to be differentiated after overlaying it with the CBCACONH spectrum. Theoretically, complete backbone assignment is possible using this pair of experiments alone. However, overlapping of peaks and the inherently low sensitivity of these experiments (see Table 6.2) means that further triple resonance experiments need to be performed.

The HNCA/HNCOCA and HNCO/HNCACO pairs of experiments are frequently employed to aid in backbone assignment (Bax and Ikura, 1991; Clubb et al., 1992; Kay et al., 1990). These experiments are much more sensitive and can resolve several ambiguities encountered with the CBCACONH/CBCANH pair. Furthermore, these experiments can be performed with a better resolution in the ^{13}C dimension due to the fact that the sweep width is smaller due to a smaller chemical shift range (compared with the CBCACONH/CBCANH pair) and are extremely useful in forming amino acid links. Figures 1.14(b) & (c) show the magnetisation transfer pathways for the HNCA/HNCOCA and HNCO/HNCACO pairs of experiments respectively.

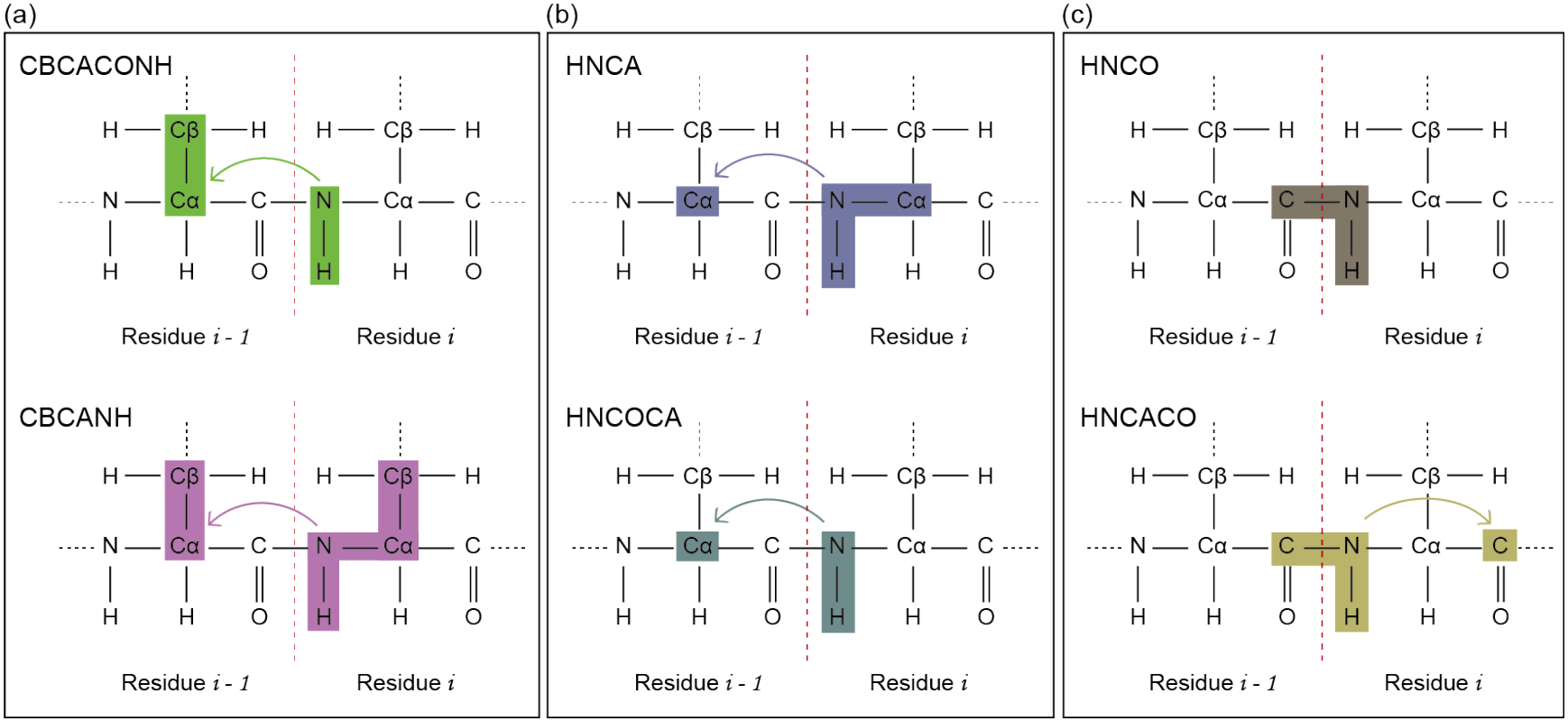


Figure 1.14: Correlations for triple resonance experiments. (a) During CBCACONH, the residue (i) is correlated with Cα and Cβ of the preceding residue ($i-1$) whereas in CBCANH the residue (i) is correlated with Cα and Cβ of the same residue as well as the previous residue; (b) HNCA correlates the residue (i) with the Cα of the same residue as well as the previous residue ($i-1$) whereas HNCOCA correlates the residue (i) with Cα of the previous residue ($i-1$); (c) HNCO correlates residue (i) with the CO of the preceding residue whereas HNCACO provides the same correlation in addition correlating it the CO of the current residue.

1.6 Solid-State NMR Theory

The chemical shift of a nucleus changes with respect to its orientation in a magnetic field. This orientation dependence is known as anisotropy.

1.6.1 Anisotropic interactions

While anisotropic interactions are averaged out in solution-state NMR due to molecular tumbling, solid samples possess restricted molecular motion and the full effect of orientation-dependent interactions is still present in solid-state NMR. As molecules are oriented in possible directions with respect to the external magnetic field, superposition of all possible chemical shifts is observed for each nucleus giving rise to broad peaks.

Two of the most important anisotropic interactions in solid-state NMR are chemical shift anisotropy (CSA) and dipolar coupling. Figure 1.15 shows how the CSA of a ^{13}CO nucleus with respect to a static magnetic field (B_0) affects the resonance frequency. The resultant spectrum shows a sum over distribution of orientations of the nucleus with respect to the magnetic field.

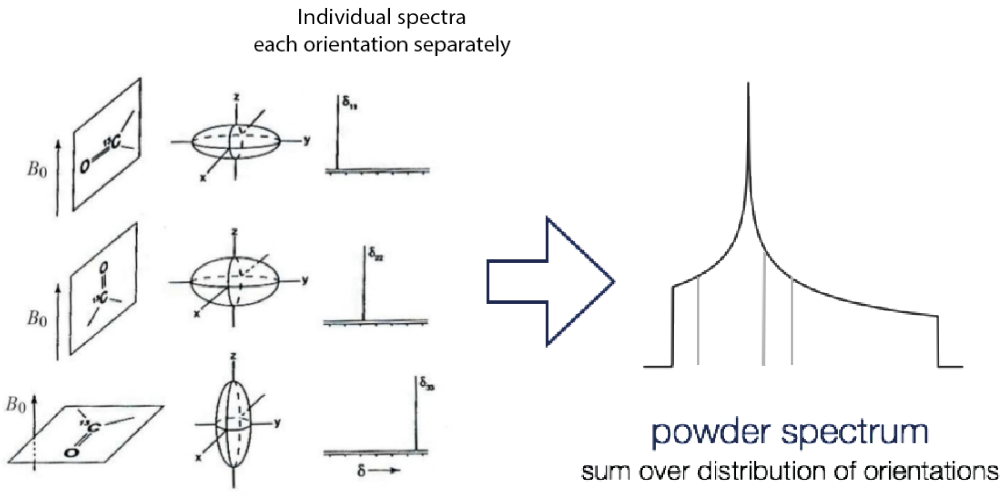


Figure 1.15: Dependence of resonance frequency upon orientation of ^{13}CO due to chemical shift anisotropy. Adapted from (Gauglitz and Moore, 2014).

Similarly, dipolar coupling, which arises as a result of the nuclear interaction between

two separate spins, is also a cause for line broadening. It is dependent on the orientation of the vector connecting the two nuclear spins.

Although both of the aforementioned anisotropic interactions significantly hinder resolution of distinct sites, they still contain valuable structural and dynamic information about the protein. NMR spectroscopists have developed a number of ingenious approaches to access this information. One such approach is to mimic the averaging effect in solution-state NMR due to tumbling using a technique called magic angle spinning (MAS).

1.6.2 Magic angle spinning (MAS) NMR

First introduced by E.R. Andrew and I.J. Lowe (Andrew et al., 1959), magic angle spinning is a routinely used technique in solid-state NMR to mimic isotropic tumbling in the solution state by the mechanical rotation oft he sample. The reason why magic angle spinning is successful is that anisotropic interactions such as CSA and dipolar coupling have the same orientational dependence:

$$\tfrac{1}{2}(3\cos^2\theta - 1)$$

where θ denotes the angle between anisotropic interaction vector (CSA or dipolar coupling) and the magnetic field B_0. Thus, for all vectors where $\theta = 54.7^\circ$ the anisotropic component of the relevant interaction is zero.

When a sample is placed in a magnetic field and spun at high frequency about an axis 54.7° to the magnetic field, it simulates isotropic tumbling. This can be understood by considering a component of the interaction vector parallel to the magic angle and perpendicular to the magic angle. All vectors parallel to the rotation axis will be zero on average as they are at the magic angle. And all vectors perpendicular to the rotation axis will be zero on average due to high frequency spinning.

1.6.3 Cross polarisation (CP)

Cross polarisation (CP) is a central theme in solid-state NMR and involves transferring magnetic polarisation from an abundant nucleus with a high gyromagnetic ratio (usually ^{1}H) to a nucleus with a low gyromagnetic ratio such as ^{13}C. The primary consequence of this is an enhancement in sensitivity. CP is achieved by RF irradiating simultaneously both resonance frequencies such that the transition energy for both nuclei is the same. This is also known as the Hartmann-Hahn matching condition.

CP usually provides a much higher signal to noise ratio than direct excitation, and allows a faster repetition rate for the experiments (it now depends on the T_1 of ^{1}H). However, since it requires high-power irradiation, it could have deleterious effects on the sample or probe and these must be monitored closely.

1.6.4 Dipolar assisted rotational recoupling (DARR) experiment

Dipolar Assisted Rotational Recoupling (DARR) (Takegoshi et al., 2001) is one of the most popular ^{13}C–^{13}C correlation experiments in solid-state NMR and the method of choice to observe such correlations in this study. While MAS successfully averages out anisotropic interactions that could lead to spectra which are unresolvable, it also deprives us of a wealth of structural information regarding orientation and inter-nuclear distances. In light of this, several RF pulse sequences have been developed to selectively reintroduce desired interactions. DARR is one such mechanism which is designed to recouple homonuclear dipolar couplings under MAS. DARR has been, and continues to be, applied to a number of membrane protein systems to observe short- and long-range through space ^{13}C–^{13}C correlations (Shahid et al., 2012b; Das et al., 2013; Yamamoto et al., 2015). Figure 1.16 shows the pulse sequence for the DARR experiment while the magnetisation transfer between nuclei during the DARR regime is illustrated in Figure 1.17.

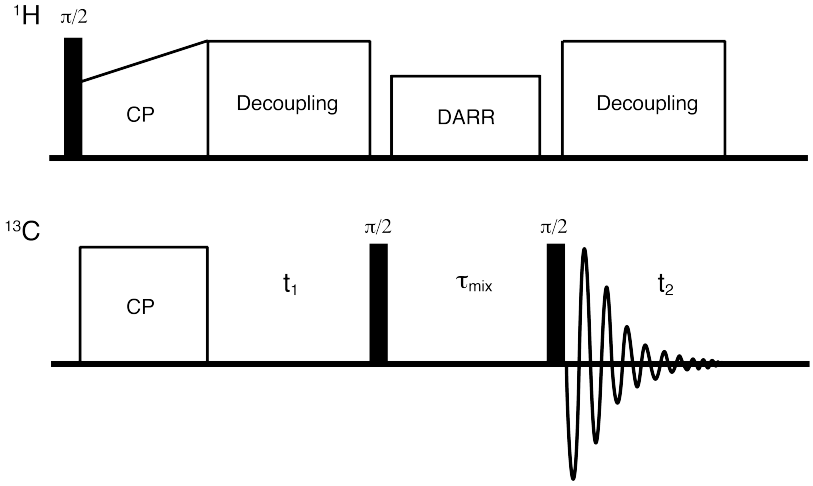


Figure 1.16: Pulse sequence for Dipolar assisted rotational recoupling (DARR) experiment. The ^{1}H magnetisation is first transferred to the xy-plane through a $\frac{\pi}{2}$-pulse before applying a ramped CP pulse to transfer polarisation to ^{13}C. ^{1}H is decoupled during the t_1 evolution after which ^{13}C magnetisation is placed on the z-axis using a $\frac{\pi}{2}$-pulse. A lower power DARR pulse on the ^{1}H channel equal to that of the MAS spinning speed is also applied. A final $\frac{\pi}{2}$-pulse is applied on the ^{13}C before recording the FID, during which time a high power decoupling pulse is applied on the ^{1}H channel.

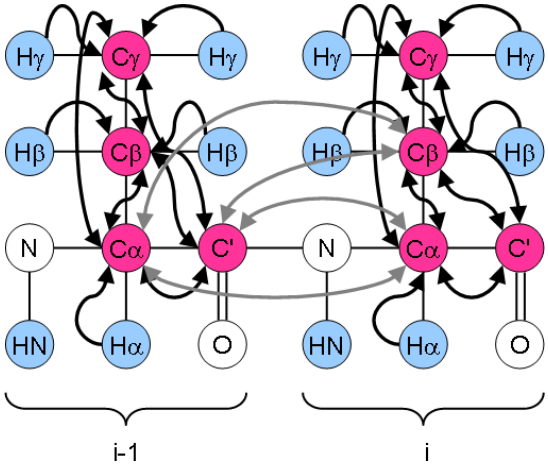


Figure 1.17: Magnetisation transfer in a protein during a DARR experiment. Nuclei circled in pink indicate those atoms observed in the experiment and those circled in blue indicate those atoms through which magnetisation flows in a protein chain. Residues (i) and $(i-1)$ are used to illustrate this. Black arrows show magnetisation transfer and grey arrows indicate magnetisation that is normally only observed through longer mixing times (image taken from `http://www.protein-nmr.org.uk/spectra_ssnmr.html`).

1.7 Aims and Objectives

As has been highlighted in section 1.2, the malfunction of RTKs has been linked to a plethora of diseases. The transmembrane domain, albeit a short span in the sequence of these RTKs, is of critical importance as a single mutation can lead to the constitutive activation of a receptor. This study aimed at using solution- and solid-state NMR to probe the structure and interactions of two biologically important RTK-TMs for which no high resolution structures exist.

The first RTK-TM to be studied was oncogenic Neu. Structural insights into what amino acid motifs stabilise the activated dimer of oncogenic Neu were sought by selectively labelling amino acids that have been proposed to lie at the dimer interface. After reconstituting them into 'native-like' lipid bilayers, through-space NMR (DARR) correlation experiments were performed to identify which amino acids lie close in space in the dimeric interface of Neu.

An alternative approach was chosen to study PDGFβR-TM. The protein was isotopically labelled so as to be able to assign the protein backbone. Once this was done, E5 was titrated into the sample to obtain not only direct evidence of binding, but to also deduce exactly which amino acids are involved in stabilising the PDGFβR-TM/E5 complex. In doing so, a detailed set of protocols for the cloning, expression, purification and reconstitution of PDGFβR-TM were developed.

Successful development of these methods and their application to RTKs will provide critical information about their behaviour in various diseases and aid the development of rationally-designed drugs. In pursuit of this long term objective of structural biology, this study focuses on the transmembrane domains of oncogenic Neu and PDGFβR.

Chapter 2

Materials and Methods

2.1 Suppliers of Chemicals and Reagents

All chemicals and reagents used in this study were of an appropriate purity grade (e.g. analytical or molecular biology) and were acquired from the following suppliers unless otherwise specified:

- Fisher Scientific UK Ltd.; Loughborough, UK
- Promega Corporation; Madison, Wisconsin, USA
- Sigma-Aldrich Company Ltd.; Dorset, UK
- New England Biolabs (UK) Ltd.; Hitchin, Hertfordshire, UK
- Perkin Elmer; Waltham, Massachusetts, USA
- Affymetrix (formerly Anatrace); Santa Clara, California, USA
- GE Healthcare; Little Chalfont, Buckinghamshire, UK
- Avanti Polar Lipids; Alabaster, Alabama, USA
- VWR International Ltd.; Lutterworth, UK
- Roche; Basel, Switzerland
- Life Technologies Ltd.; Paisley, UK

2.2 Plasmids

The pET151/D-TOPO plasmid was purchased directly from Life Technologies UK. The pDEST17-MISTIC plasmid was kindly provided by authors of (Bernaudat et al., 2011b) at the Institut de Biologie Structurale Jean-Pierre Ebel, CEA, Grenoble, France. Previously generated plasmids used in this study and the plasmids generated in this study are illustrated in Table 2.1 and Table 2.2 respectively.

Plasmid	Details	Reference
pET30a-PDGFR20	pET30a vector ligated with PDGFβR transmembrane domain and 20 residues on either side beteeen BamH1 and Xho1 restriction sites	(Oates et al., 2010)
pET30a-PDGFR30	pET30a vector ligated with PDGFβR transmembrane domain and 30 residues on either side beteeen BamH1 and Xho1 restriction sites	(Oates et al., 2010)
pET30a-PDGFR40	pET30a vector ligated with PDGFβR transmembrane domain and 40 residues on either side beteeen BamH1 and Xho1 restriction sites	(Oates et al., 2010)
pET151/D-TOPO	An expression vector to be used for directional TOPO cloning with a hexahistidine tag and TEV-specific proteolysis site	(Shuman, 1994)
pDEST17-MISTIC	A Gateway cloning vector with the MISTIC-tag added	(Bernaudat et al., 2011b)
pET21a-GFP-TEV	A vector for the expression of GFP-TEV protease with a His-tag	(Wu et al., 2009)
pRK793-TEV	A vector for the expression of MBP-TEV protease with a His-tag	(Kapust et al., 2001)

Table 2.1: Previously generated plasmids used in this study

Plasmid	Template	Features
pET151-PDGFR20	pET151/D-TOPO	Gene of PDGFβR-TM plus 20 residues either side inserted in the TOPO cloning site
pDEST17-MISTIC-PDGFR	pDEST17-MISTIC	Gene of PDGFβR-TM plus 20 residues either side inserted between PspXI and SalI restriction sites

Table 2.2: Plasmids generated in this study

2.3 DNA Manipulation and Cloning Techniques

2.3.1 Primers

Primers were designed against the relevant gene target to allow the cloning of PCR products in frame with any C-terminal tags. Restriction sites or overhang sequences for TOPO-cloning were also incorporated, as appropriate. All primers were purchased from Sigma-Aldrich, UK.

Name	Sequence 5' to 3'	Tm (°C)
topoforward	GTGGAACTCCATGGGTGGAGATTCGCAGG	80
toporeverse	CTATTAACCGTCAGAGCTCACAGACTCAATGA	71.7
pdgfrforward	TAAGCACATATGATGTCGTACTACCATCACCATC ACCATCAC	75.5
pdgfrreverse	TCCGATGCTCGAGGCTATTAACCGTCAGAGCTC ACAGACTCAAT	80.9

Table 2.3: Primers used in this study, Tm represents the melting temperature

2.3.2 Preparation and storage of plasmid DNA

Plasmid DNA was isolated using GeneJET MiniPrep kit (Thermo Fisher Scientifc, UK) as per manufacturer's protocol. The DNA was recovered in 50 μL dH_2O and stored at -20 °C.

2.3.3 Polymerase chain reaction (PCR)

The amplification of the PDGFβR-TM gene was performed on ice with 5 μL of 5× Phusion HF buffer, 1 μL 10 mM dNTP mix, 2.5 μM of each primer, 1 μL template DNA (0.5 ng), 0.5 μL of Phusion DNA Polymerase (1 units/50 μL) and dH_2O to a final volume of 50 μL.

After gently mixing the reaction, all liquid was collected at the bottom of the PCR tube through a quick spin in a microcentrifuge. The tubes were then transferred

from ice to a Biometra T3 thermocycler and a PCR was initiated according to the conditions illustrated in Table 2.4.

Step	Temperature (°C)	Time (s)
Initial Denaturation	98	30
25-35 Cycles	98	10
	$T_{annealing}$	30
	72	20 per kb
Final Extension	72	600
Hold	4	

Table 2.4: Thermal cycling scheme for a typical PCR, $T_{annealing}$ was 3–5 °C below the Tm of the relevant primers

2.3.4 Agarose gel electrophoresis

Agarose gels were prepared by dissolving 1–2% high-melting point agarose (Sigma-Aldrich) in 50 mL of 1× Tris-acetate-EDTA (TAE) buffer in a 250 mL conical flask and heating in a microwave oven until dissolved (∼2 minutes). The solution was subsequently cooled and 5 μL of Biotium GelRed Nucleic Acid Stain (10,000× in water) was added and gently shaken to mix evenly. This was immediately poured into a gel cast and left to solidify. DNA samples were prepared by adding 1 μL of 6×DNA loading dye (Fermentas) per 5 μL of DNA sample. The samples were then loaded on to a gel which had been submerged in a gel tank containing 1× TAE buffer. In order to quantify the size of the DNA samples, a lane with 10 μL of FastRuler^{TM}Middle Range DNA ladder was added. Electrophoresis was performed at 150 V for 30 minutes or more, as required. The DNA was visualised using BioDoc-It® Imaging System equipped with a camera to take photographs of the gels with the DNA illuminated.

2.3.5 Purification of DNA from PCR and restriction digests

Purification of the DNA samples from polymerases and endonucleases after PCR and restriction digests was performed through adsorption to a silica matrix using QIAquick PCR Purification Kit (QIAGEN). In cases where gel extraction was re-

quired, the QIAquick Gel Extraction Kit (QIAGEN) was used which also achieves purification through the adsorption of DNA on a silica matrix. Both kits were used as per manufacturer's instruction and DNA was eluted in 30 μL of water.

2.3.6 Restriction endonuclease digestion of DNA

Restriction digests were carried out in an appropriate buffer system using restriction enzymes (NEB, UK) as per the manufacturer's instructions. Typically, reactions were incubated for 2–3 hours at a temperature of 37 °C. DNA was subsequently purified as detailed in section 2.3.5.

2.3.7 Ligation of chimeric DNA fragments

After gel extraction and/or purification, the linearised vector strands and insert DNA fragments were ligated using T4 DNA Polymerase (NEB, UK) as per manufacturer's protocol. This was typically done for 2 hours at room temperature or overnight at 4 °C at a vector to insert ratio of 1 : 3.

In particularly troublesome ligation reactions, calf-intestinal alkaline phosphatase (CIP) (NEB, UK) was used, as per manufacturer's instructions, to dephosphorylate the 5'-ends of linearised vector DNA to prevent recircularisation.

The ligation mixture was then used to directly transform the relevant strain of *E. coli* as per the description in section 2.5.3. Colonies were picked at random using a sterile pipette tip and cultured overnight at 37 °C in 10 mL of LB in the presence of appropriate antibiotics. The plasmid DNA was then extracted (section 2.3.5) and digested using restriction enzymes (section 2.3.5). It was then analysed on an agarose gel (section 2.3.4) to establish the presence of an insert at the appropriate molecular weight.

2.3.8 Colony PCR

In order to affirm the presence of insert DNA in the plasmid constructs, a colony PCR was performed. This was done by picking a random selection of 30 colonies and resuspending each of the them in 50 μL of PCR mixture described in section 2.3.3. The PCR was run according to the relevant thermal profile and the reactions were subsequently run on agarose gels. Where there was evidence of the insert being present, the colony was grown overnight in rich media after which the plasmid was extracted and sent for sequencing.

2.3.9 Concentration of DNA fragments

The concentration of DNA fragments was measured using Thermo Scientific NanoDrop Lite spectrophotometer.

2.3.10 DNA sequencing of plasmid constructs

For all cloned genes, it was important to ensure that

- start and stop codons were present,
- there were no PCR errors in the form of point mutations etc.,
- the gene along with any tags or restriction sites that were added was correctly oriented and in frame.

In order to do so, 20 μL of $\sim$100 ng/μL of DNA was sent for sequencing to GATC Biotech Ltd., UK. All sequencing experiments were performed using a T7 primer. Sequencing results were manually checked using Serial Cloner 2.6 for Mac OS X.

2.4 *Escherichia coli* Strains

The *E. coli* strains used in this study are illustrated in Table 2.5.

Strain	Genotype	Reference
DH5α	F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA - *arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk-,mk+) *phoA sup*E44 λ-*thi*-1 *gyr*A96 *rel*A1	(Hanahan, 1983)
TOP10	F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	(Grant et al., 1990)
BL21(DE3)	*fhuA2 [lon] ompT gal* (λ *DE3) [dcm]* Δ *hsdS* λ *DE3* = λ *sBamHIo* Δ *EcoRI-B int:: (lacI:: PlacUV5:: T7 gene1) i21* Δ *nin5*	(Studier and Moffatt, 1986)
BL21(DE3)pLysS	F$^-$ *ompT hsdSB (rB- mB-) gal dcm* (DE3) pLysS Cm^r	(Moffatt and Studier, 1987)
C41(DE3)	F$^-$ *ompT hsdSB (rB- mB-) gal dcm* (DE3)	(Miroux and Walker, 1996a)
C43(DE3)	F$^-$ *ompT hsdSB (rB- mB-) gal dcm* (DE3)	(Miroux and Walker, 1996a)
BL21-CodonPlus (DE3)-RIPL	F$^-$ *ompT hsdSB (rB- mB-) dcm+* Tet^r *gal* λ (DE3) *endA* Hte [*argU proL* Cam^r] [*argU ileY leuW* Strep/Spec^r]	(Jerpseth et al., 1998)
Rosetta^{TM}(DE3) pLysS	F$^-$ *ompT hsdSB (rB- mB-) gal dcm* (DE3) pLysSRARE (Cam^R)	(Jerpseth et al., 1998)
ccdB Survival^{TM}	F$^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G *fhu*A::*IS*2	(Bernard and Couturier, 1992; Bernard et al., 1993)

Table 2.5: Bacterial strains used in this study

2.5 Growth and Maintenance of *E. coli*

2.5.1 Growth Media

Lysogeny broth (LB)

Lysogeny Broth (Bertani, 2004) is a rich growth medium and is composed of 10 g/L bacto-tryptone, 5 g/L yeast extract, 10 g/L sodium chloride. It is commonly used for the aerobic growth of *E. coli* in liquid media. LB was prepared in deionised water (dH_2O) and autoclaved prior to use.

In order to make LB agar plates, 15 g of bacto-agar was added to 1 L of media and autoclaved. This solution was then cooled to ~50 °C and poured onto sterile petri

dishes after the addition of appropriate antibiotics. These were then sealed using paraffin film and stored at 4 °C.

Super optimal broth with catabolites (SOC) medium

This was originally proposed by Hanahan and coworkers (Hanahan, 1983) and is composed of 20 g/L tryptone, 0.5 g/L NaCl, 5 g/L yeast extract, 0.2 g/L KCl, 3.6 g/L glucose and 1 g/L $MgCl_2$. This nutrient-rich medium was used for outgrowths after bacterial transformation as it is known to result in higher transformation efficiency. SOC medium was prepared in dH_2O and autoclaved prior to use.

M9 minimal medium

5× M9 salts were prepared by dissolving 64 g $Na_2HPO_4.7H_2O$, 15 g KH_2PO_4 and 2.5 g NaCl in dH_2O and autoclaving. Subsequently, 1 L of the medium was prepared by adding 100 μL of 1 M $CaCl_2$, 2 mL of 1 M $MgSO_2$, 1 mL of 15 mg/mL $FeCl_3$ and 10 mL of 100× BME vitamins (Sigma) in ∼500 mL of dH_2O. 1 mg of NH_4Cl (or ^{15}N-NH_4Cl) and 4 mg of glucose (or ^{13}C-glucose or ^{13}C-1-glucose) were then added and completely dissolved. Finally, appropriate antibiotics were added and the pH was adjusted to 7.4 before topping up to 1 L with dH_2O.

Antibiotics were added to all media, where necessary, to concentrations in Table 2.6.

Antibiotic	Concentration μg/mL
Ampicillin	100
Kanamycin	50
Chloremphenicol	35
Carbenecillin	100

Table 2.6: Antibiotic concentrations

Handling of the aforementioned media was done within a sterile field created by a Bunsen burner flame until the medium was sealed.

2.5.2 Preparation of competent *E. coli* cells

To generate chemically competent *E. coli* cells, 10 mL of LB was inoculated with the relevant strain of *E. coli* cells and incubated overnight at 37 °C with shaking at 180 rpm. Subsequently, 2.5 mL of this culture was added to 250 mL of fresh LB and grown under the same conditions until an optical density at a wavelength of 600 nm (OD_{600}) of 0.4–0.6 was reached. The cells were then pelleted by centrifugation at 3000 × *g* and resuspended in 10 mL of 50 mM $CaCl_2$. After leaving this suspension on ice for 5 minutes, the centrifugation was repeated and the cells were once again resuspended in 100 mL of 100 mM $CaCl_2$ plus 15% v/v glycerol. This was then incubated on ice for ∼1 hour before being snap frozen in liquid nitrogen to be stored as 100 μL aliquots at -80 °C. All solutions were filter-sterilised prior to use.

2.5.3 Bacterial transformations

Competent cells, whether provided by a commercial supplier or prepared in the laboratory as described in section 2.3.5, were first thawed on ice. A 1–2 μL aliquot of DNA was added to 50 μL of competent cells and mixed gently followed by incubation on ice for 30 minutes. Cells were then heat-shocked at 42 °C for 45 seconds and then placed on ice for 2–3 minutes. Following this, 1 mL of pre-warmed LB or SOC medium was added to the cells and the suspension was incubated at 37 °C with shaking at 180 rpm for 1 hour. The cells were then pelleted by centrifugation at 3000 *g* for 5 minutes. A total of 850 μL of clear LB was removed from the supernatant and the pelleted cells were resuspended in the remaining 200 μL. This was then plated on LB-agar plates containing the necessary antibiotic selection for the transformed plasmid. Plates were incubated at 37 °C overnight.

2.5.4 Preparation of glycerol stocks

Plasmid constructs which demonstrated long-term stability in bacterial cells were stored after transformation as glycerol stocks. These were made by inoculating 5 mL

of LB supplemented with the appropriate antibiotic(s) with a single *E. coli* colony following a fresh transformation. The cells were grown at 37 °C with shaking at 180 rpm to mid-log phase and then mixed with sterile 50% v/v glycerol in a 1 : 1 ratio in a Corning cryo-vial. This was then frozen in liquid N_2 and stored in a -80 °C refrigerator until required.

2.6 PDGFβR-TM Expression and Purification

2.6.1 Expression of platelet-derived growth factor receptor-β transmembrane domain (PDGFβR-TM)

PDGFβR-TM was heterologously expressed with 20, 30 or 40 residues of the receptor on either side. These peptides were expressed in various *E. coli* strains (Table 2.5) with pET expression systems using IPTG induction. Their expression was optimised in rich and minimal media so as to obtain high enough yields for biophysical analyses. Competent cells of the relevant strains were transformed with an expression plasmid containing the PDGFβR-TM gene and plated on LB agar plates with appropriate antibiotics.

Overexpression of the target gene was carried out either in a rich (LB) or minimal (M9) medium. However, in all cases, after fresh transformation of the cells (section 2.5.3), a colony was picked at random and grown overnight in 10 mL of LB at 37 °C with shaking at 180 rpm in the presence of appropriate antibiotics. This pre-culture was then added to 1 L of fresh LB containing the same antibiotics. When the pET151/D-TOPO vector was being used, the pre-culture was centrifuged and then resuspended in 1 L of LB. The cells were then incubated in a 5 L conical flask under the same conditions until and OD_{600} of 0.6 was reached. The expression protocol differed henceforth depending on the expression medium.

Expression in LB

In order to induce expression, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the culture to a final concentration of 1 mM. This metabolite initiates the transcription of the *lac* operon hence triggering production of the protein of interest. The cells were harvested after an additional 4–5 hours of incubation using SORVALL® RC 6 Plus centrifuge at 3000 g for 10 minutes. They were washed with Tris-HCl buffer (pH 7.5) and stored at -80 °C until required.

Expression in M9 minimal medium

Following a protocol for the production of very high yields of recombinant proteins (Sivashanmugam et al., 2009), just prior to induction, the cells were harvested and washed with 1× M9 salts to remove any residual LB. The pellet was then resuspended in a smaller volume of M9 minimal media (section 2.5.1) using fraction (1/2, 1/3, 1/4 or 1/8) of the initial volume. The cells were then left to equilibrate in M9 media (37 °C with shaking at 180 rpm) for an hour before being induced by 1 mM IPTG and left to overexpress for 4–5 hours. They were harvested, washed and stored in the same way as described above.

Aliquots were taken pre-induction and post-expression for SDS-PAGE (section 2.8.1) and Western Blot analysis (section 2.8.4) for all PDGFβR-TM expressions.

2.6.2 Purification of PDGFβR-TM for solid-state NMR

The cells from 1 L of culture were thawed on ice and then resuspended in 20 mL of lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl). Additionally, 20 μL of 10 mg/mL DNAse I was added along with $MgCl_2$ to a final concentration of 5 mM. Lysozyme was also added to a final concentration of 1 mg/mL. This was left to incubate on ice for 30 minutes. The cells were then lysed either using a probe sonication (30 second pulses with 30 second gaps for 10 minutes) or using a Constant Cell Disruption System (10 mL × 2). The inclusion bodies along with cell debris

were then pelleted using a SORVALL® RC 6 Plus centrifuge with an SS-34 rotor for 90 minutes at 47,000 × *g*. The pellet was then washed with 5 mL lysis buffer and then centrifuged for an additional 20 minutes. The supernatant was then stored at 4 °C and the pellet was dissolved in 20 mL solubilisation buffer (100 mM NaCl, 50 mM NaH_2PO_4, 6 M GnHCl (or 8 M urea), pH 8.0) while rotating on a end-over-end mixer at 4 °C overnight.

After the inclusion bodies were solubilised, the mixture was centrifuged again at 47,000 × *g* for 30 minutes. The insoluble material was stored and the supernatant was passed through a 4.5 μm filter. The purification procedure was then continued with immobilised metal ion affinity chromatography (IMAC) (see section 2.6.4 for more details). The solution was poured in a Bio-Rad Econo-Column® chromatography column and mixed with ∼4 mL of IMAC Sepharose High Performance resin (GE Healthcare) charged with Ni^{2+} ions and left to rotate on an end-over-end mixer for ∼1 hour. The resin should only bind His-tagged proteins due to the presence of charged metal groups on the beads. The suspension was then left to settle and the flow-through was collected. The resin was then washed with 20 mL of wash buffer (20 mM HEPES, 150 mM NaCl, 5% glycerol, 1 mM n-decyl-β-D-maltopyranoside (or 1% Triton X-100), 20 mM imidazole, pH 8.0). The presence of detergent also ensured on-column refolding of the PDGFβR-TM peptide, whereas glycerol was used to stabilise the protein by reducing hydrophobic interactions. The protein was then eluted from the column by washing the beads with 2.5 mL elution buffer (20 mM HEPES, 150 mM NaCl, 5% glycerol, 1 mM n-decyl-β-D-maltopyranoside (or 1% Triton X-100), 330 mM imidazole, pH 8.0). This was repeated twice to ensure all the protein had eluted off the column. The solution was then immediately passed through a PD10 desalting column (GE Healthcare) to remove imidazole, as per manufacturer's instructions.

All buffers used in the purification process were either filter-sterilised or autoclaved, as appropriate. Aliquots were taken for SDS-PAGE (section 2.8.1) and Western Blot analysis (section 2.8.4) for each step in the purification.

2.6.3 Purification of PDGFβR-TM for solution-state NMR

The cells from 1 L of culture were thawed on ice and then resuspended in 20 mL lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl). This was followed by an addition of 2 mL of Novagen BugBuster^{TM}Protein Extraction Reagent and this left to incubate at room temperature for about an hour with gentle shaking. The cells were then lysed using a Constant Cell Disruption System (10 mL $\times$ 2). The inclusion bodies along with cell debris were then pelleted using a SORVALL® RC 6 Plus centrifuge with an SS-34 rotor for 45 minutes at 47,000 $\times$ *g*. The supernatant was then stored at 4 °C and the pellet was dissolved in 20 mL solubilisation buffer (100 mM NaCl, 20 mM Tris-HCl, 6 M GnHCl (or 8 M urea), pH 8.0) while rotating on a end-over-end mixer at 4 °C overnight.

After the inclusion bodies were solubilised, the mixture was centrifuged again at 47,000 $\times$ *g* for 30 minutes. The insoluble material was stored and the supernatant was passed through a 0.45 μm filter. The purification procedure was then continued with IMAC (2.6.4). The solution was poured in a Bio-Rad Econo-Column® chromatography column and mixed with ∼4 mL of IMAC Sepharose High Performance resin (GE Healthcare, UK) charged with Ni^{2+} ions and left to rotate on an end-over-end mixer for ∼1 hour. The suspension was then left to settle and the flow-through was collected. The resin was then washed with 20 mL of wash buffer (100 mM NaCl, 20 mM Tris-HCl, 6 M GnHCl (or 8 M urea), 10 mM imidazole, pH 8.0). The protein was then eluted from the column by washing the beads with 2.5 mL elution buffer (100 mM NaCl, 20 mM Tris-HCl, 6 M GnHCl (or 8 M urea), 10 mM imidazole, pH 8.0). The solution obtained was then pooled and dialysed in a dialysis tubing with 3.5 kDa cut-off against 4 L of water for 24 hours switching to fresh water at regular intervals. Once all the protein had precipitated out, it was pelleted and freeze-dried to obtain a protein powder.

It was found that it is advantageous to further purify this peptide using reverse-phase high performance liquid chromatography (HPLC). To that end, approximately 2 mg of this protein powder was dissolved in 200 μL of trifluoroacetic acid (TFA)

and 200 μL of 2,2,2-trifluoroethanol (TFE) and the suspension was sonicated until complete dissolution of the powder was observed. 400 μL of isopropanol was then added followed by the gradual addition of $\sim$1 mL of a solution containing HPLC-grade water and isopropanol in a 3 : 7 ratio. This mixture was then loaded on to a semi-preparative Jupiter C4 5 μm (300 Å, 250 $\times$ 10.0 mm) reverse-phase HPLC column (Phenomenex, Macclesfield, Cheshire, UK) connected to a purpose built two-pump HPLC system (Jasco UK, Great Dunmow, Essex, UK). The column was equilibrated with 70% Buffer A and 30% Buffer B (see below for compositions). While maintaining a constant flow rate of 1.5 mL/minute, the protein was purified using a two-component gradient illustrated in Table 2.7. The two HPLC buffers were the same as those used by (Claridge and Schnell, 2012):

- Buffer A: 5% isopropanol, 95% water, 0.1% trifluoroacetic acid
- Buffer B: 57% isopropanol, 38% acetonitrile, 0.1% trifluoroacetic acid

Time (minutes)	Buffer A (%)	Buffer B (%)
0	70	30
10	70	30
20	0	100
30	0	100
40	70	30
50	70	30

Table 2.7: HPLC gradient for the purification of PDGFβR-TM

The presence and purity of PDGFβR-TM in a particular fraction was confirmed through mass spectrometry (section 2.8.5). All the fractions containing the peptide were then pooled and freeze-dried to obtain PDGFβR-TM in powder form. This was then stored at 4 °C until required.

2.6.4 Immobilised-metal affinity chromatography (IMAC)

IMAC was first introduced and shown as a feasible technique for protein separation by Porath and coworkers (Porath et al., 1976). It exploits the fact that transition metal ions (such as Ni^{2+} and Co^{2+}) have an affinity for histidine and cysteine residues in solution. This technique was further improved upon the introduction of

nitrilotriacetic acid (NTA) as an extremely effective chelating ligand (Hochuli et al., 1987).

In this study, all the constructs used to express PDGFR-TM domains and TEV protease encoded for an N-terminal hexahistidine tag to allow IMAC purification (sections 2.6.3 & 2.7.2). The protein was resolved following a gravity-flow protocol with an IMAC Sepharose High Performance resin (GE Healthcare, UK) packed in a Bio-Rad Econo-Column® chromatography column. The resin was first washed with dH_2O and if the resin was previously used, it was stripped with 5 column volumes of stripping buffer buffer (20 mM Tris-HCl, 50 mM EDTA, 200 mM NaCl, pH 7.4). It was then washed with 10 column volumes of dH_2O. The column was then charged with 2 column volumes of $NiCl_2$ solution and washed again with 10 column volumes of dH_2O and stored in 20% ethanol until required.

2.6.5 Plasmid stability test

It is advisable to check the fraction of cells that still carry the plasmid just prior to induction. This can be used to maximise protein yield by optimising variables in the protein expression protocol such as induction OD_{600}, bacterial cell line and choice of antibiotics. This involves preparing four different plates (see Table 2.8).

Plate	Cells that grow
LB	All viable cells
LB + antibiotic	Cells that still carry the plasmid
LB + IPTG (1 mM)	Cells that have lost the plasmid or mutants that have lost the ability to express the target gene
LB + antibiotic + IPTG (1 mM)	Mutants that retain the plasmid but have lost the ability to express the target gene

Table 2.8: Plasmid stability test

An aliquot of cell culture is taken immediately before induction. A serial dilution of this is performed including a 10^5 and a 10^6 dilution. Cells from the 10^5 dilution are plated on the LB + IPTG plate and on the LB + IPTG + antibiotic plate. Cells from the 10^6 dilution are plated on the LB plate and on the LB + antibiotic plate. These are incubated overnight at 37 °C.

Cells do not grow on the IPTG plates if they are carrying a protein production plasmid because they have dedicated all their resources to the production of the recombinant protein instead of cell proliferation. With unstable target plasmids, the fraction of cells that have lost the plasmid will be reflected by an increase in colonies on the LB + IPTG plate and a decrease on the LB + antibiotic plate.

2.7 Tobacco etch virus (TEV) protease expression and purification

The construct for MBP-tagged TEV protease was transformed into BL21-CodonPlus (DE3)-RIPL cells whereas the GFP-tagged TEV protease was transformed into Rosetta^{TM}(DE3) pLysS as described in section 2.5.3. The protocol adopted for the expression and purification of maltose binding protein (MBP)- or green fluorescent protein (GFP)-tagged TEV protease were the same.

2.7.1 Expression

A 10 mL starter culture grown overnight at 37 °Ccontaining appropriate antibiotics was inoculated from glycerol stocks (see section 2.5.4) using a sterile loop. This starter culture was then poured into 1 L of LB and grown at 37 °C with shaking at 180 rpm in the presence of appropriate antibiotics until an OD_{600} of 0.6 was reached. The cells were then induced using 0.1 mM IPTG and incubated for 24 hours at 15 °C with shaking at 180 rpm. The cells were then harvested using SORVALL® RC 6 Plus centrifuge at 3000 × g for 10 minutes. They were washed with Tris-HCl buffer (pH 7.5) and stored at -80 °C until required.

2.7.2 Purification

The cells from 1 L of culture were thawed on ice and then resuspended in 20 mL lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl). This was followed by an addition of 2 mL of Novagen BugBuster^{TM}Protein Extraction Reagent and this left to incubate

at room temperature for about an hour with gentle shaking. The cells were then lysed using a Constant Cell Disruption System (10 mL × 2). This suspension was then centrifuged using a SORVALL® RC 6 Plus centrifuge with an SS-34 rotor for 45 minutes at 47,000 × *g*. The pellet was discarded and the supernatant was passed through a 0.45 μm filter. The purification procedure was then continued with IMAC. The following buffers were used in the subsequent purification process:

Buffer A: 50 mM Tris-HCl, 150 mM NaCl, 20 mM imidazole, pH 8.0

Buffer B: 50 mM Tris-HCl, 150 mM NaCl, 40 mM imidazole, pH 8.0

Buffer C: 50 mM Tris-HCl, 150 mM NaCl, 300 mM imidazole, pH 8.0

The sample was loaded onto a Ni-NTA column pre-equilibrated with Buffer A. It was then washed with Buffer B with at least 5 column volumes. The protein was then eluted off the column with Buffer C and immediately passed through a PD10 column to obtain TEV in 50mM Tris-HCl, 150 mM NaCl, pH 8.0. This was either used immediately or stored in 50% glycerol at -80 °C till required.

2.8 Protein Analysis and Detection

2.8.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

At various steps of the expression and purification of proteins, it was important to ascertain presence and relative amount of the protein of interest in the sample. To this end, the separation of proteins was carried out using either a Tris-glycine buffer system (Laemmli et al., 1970) or using a commercially available NuPAGE® Bis-Tris-HCl buffered system (Life Technologies, UK).

Tris-glycine gels

Typically 1.0 mm gels were prepared with a separating gel (40% of ULTRA PURE ProtoGel® (30% (w/v) acrylamide to 0.8% (w/v) bis-acrylamide); 375 mM Tris-HCl

pH 8.8; 0.1% SDS; 0.02% APS; 0.06% TEMED) and a stacking gel (13.5% ULTRA PURE ProtoGel®; 125 mM Tris-HCl pH 6.8; 0.1% SDS; 0.6% APS; 0.06% TEMED). A plastic comb was added to form the wells and the gels were left in a casting block for about 0.5 hours. The samples were prepared by mixing them with 2× NuPAGE® sample loading buffer in a 1 : 1 ratio. They were then denatured on a heating block at 60 °C for 10 minutes. Electrophoresis was carried out at room temperature with a uniform running buffer (25 mM Tris, 250 mM glycine, 0.1% (v/v) SDS, pH 8.3). A BioRad Mini-PROTEAN® Tetra Cell was used to run the samples for 1 hour at 180 V.

NuPAGE® Bis-Tris gels

The NuPAGE® 12% Bis-Tris precast gels were washed with dH_2O and docked into an XCell Sure Lock™mini-cell electrophoresis system. The samples were prepared by mixing them with 2× NuPAGE® sample loading buffer in a 1 : 1 ratio. They were then denatured on a heating block at 60 °C for 10 minutes and loaded onto the gels. After being submerged in NuPAGE® 1× MES running buffer, the samples were allowed to run down the gels at 125 mA and 200 V for 40 minutes.

All gels were loaded with either ColorPlus™Prestained Protein Marker, Broad Range (7–175kDa) or Sigma-Aldrich® Colour Marker Ultra-Low Range (1.06–26.6 kDa) for molecular weight quantification of proteins.

2.8.2 Coomassie staining

Coomassie dye binds to proteins primarily through basic amino acids (mainly arginine, lysine and histidine) under the acidic conditions of the solution (Neuhoff et al., 1988).

After the resolution of proteins by SDS-PAGE (2.8.1), the gels were placed in fixer solution (50% (v/v) ethanol, 10% (v/v) acetic acid). They were then visualised by leaving them overnight in ∼30 mL of coomassie stain solution (0.025% (w/v) coomassie blue R-250, 10% (v/v) acetic acid) with gentle shaking. The gels were

then destained using a solution containing 20% (v/v) methanol and 7% (v/v) acetic acid for approximately 4–6 hours. They were stored in dH_2O.

2.8.3 Silver staining

In instances where the peptide/protein was of low molecular weight or if the concentration was too low to achieve adequate protein visualisation through coomassie staining, silver staining was used. This method relies on the binding of silver ions to the amino acid side chains, mostly the sulfhydril and carboxyl groups (Switzer et al., 1979; Merril et al., 1981; Merril and Pratt, 1986) followed by reduction to free metallic silver (Rabilloud, 1990).

Prior to silver staining a gel, they were coomassie stained to reduce background discolouration. Subsequently, the gel was placed in fixer solution (50% acetone, 1.25% TCA, 0.015% formaldehyde) for 15 minutes. The gel was then washed with dH_2O (3× quick wash, 1× 5 minute wash, 3× quick wash). The gel was then soaked in 50% acetone for 5 minutes followed by another 5 minute incubation in 0.0167% (w/v) sodium thiosulfate solution. The gel was again quickly washed three times with dH_2O before being placed in stain solution (0.267% (w/v) silver nitrate, 0.37% (v/v) formaldehyde) for 8 minutes. The gel was again washed with dH_2O twice before being placed in developer solution (0.2 mM sodium carbonate, 0.004% (w/v) sodium thiosulfate, 0.015% formaldehyde) until protein bands appeared (typically 10–20 seconds). The gel was then quickly placed in a quencher (1% acetic acid) to stop the developing process. The gel was then washed with dH_2O and subsequently stored in it too.

2.8.4 Immunoblotting (western blotting)

To unambiguously ascertain the presence of protein and to assess the cleavage of His-tag from a particular protein, it was often necessary to analyse the separation of proteins using western blots. This method was first developed by (Towbin et al., 1979) and involves transferring the protein bands on a gel onto a nitrocellulose

membrane and subsequently using antibodies to specifically visualise certain protein or peptides.

Following SDS-PAGE (section 2.8.1), the electrophoresed proteins along with the relevant pre-stained protein ladder were blotted onto a nitrocellulose blotting membrane with 0.45 μm pore size (GE Healthcare, UK) using an XCell II^{TM}Blot Module (Life Technologies, UK) as per manufacturer's instructions. This involved equilibrating the gel in transfer buffer (100 mM Tris, 80 mM glycine, 20% methanol) before placing it on top of the nitrocellulose membrane. The transfer was then carried out in the blotting module at 35 V for 1 hour as per manufacturer's instructions.

After the transfer the membrane was blocked in 20 mL of TBST (50 mM Tris, 150 mM NaCl, 0.1% Tween 20, pH 7.6) + 10% (w/v) milk powder by gently rocking for 1 hour at room temperature. The milk was then removed and the membrane was placed in 20 mL of TBST + 10% milk powder + (1 : 3000) Sigma-Alrich® monoclonal anti-polyhistidine (mouse IgG2a isotype) antibody for 1 hour. In order to identify C-terminal His-tags, mouse histidine (C-term) tag monoclonal antibody was used (Life Technologies, UK). It was then washed twice with TBST with gentle rocking for 5 minutes. The wash buffer was then removed and TBST + 10% milk powder + (1 : 5000) anti-mouse IgG (whole molecule) alkaline phosphatase antibody (produced in goat) was used to incubate the membrane for another hour. The secondary antibody was then discarded and the membrane was washed with TBST three times for 15 minutes each. It was then left to wash for an additional 30 minutes to minimise the background in the blot. The TBST was then discarded and 3 mL of of Novex® AP Chromogenic Substrate (BCIP/NBT) was poured over the blot. Bands typically appeared after ∼30 seconds following which the membrane was washed with TBST for 5 minutes. The membrane was then left to dry before being imaged.

Dot blots

In order to determine whether a particular His-tagged protein has been expressed or is present in a solution, a convenient quick check is to perform a dot blot. This involves taking 2 μL of the relevant sample and spotting the nitrocellulose membrane with it using a narrow-mouth pipette tip. The membrane is then allowed to dry and the blot is developed as described earlier in this section.

2.8.5 Mass spectrometry

Mass spectrometry is an analytical technique that can be used to measure the mass-to-charge ratio of various charged entities within a sample. It can be used to provide information about various physical and chemical properties of a compound. These properties include masses of particles, the elemental composition of a sample or molecule, and the chemical structures of molecules, such as peptides and other chemical compounds.

Post-HPLC analysis

In this study, mass spectrometry was used after HPLC to see which fractions contained the desired peptide (of which the mass is known) and the relative amounts present in each fraction. All such characterisation was carried out on a Bruker micrOTOF. This uses an electrospray ionisation source (ESI) which allows small quantities of the peptide in solution to be taken straight from the HPLC fractions. The sample passes through a nebuliser, is dried to remove solvent and then ionised. Approximately 10 μL of a particular fraction was injected into the spectrometer while spectra were recorded in positive ion mode, measuring between 500 and 3000 m/z (mass/charge) for an average of 1.5 minutes. The spectra collected over this time were averaged and deconvoluted using DataAnalysis v.3.3 by Bruker Daltonics, UK. This deconvoluted spectrum showed the mass of the main singly charged species in the sample. The fractions were then sorted by purity and pooled before either being lyophilised for storage, or reconstituted in membrane mimetics.

Post-SDS-PAGE analysis

Coomassie stained gel pieces were sent to the Warwick Proteomics Facility (University of Warwick), where they underwent tryptic digestion as per manufacturers protocol on the MassPrep robotic protein handling system. The extracted peptides from each band were processed through a nanoLC-ESI-MS/MS using the NanoAcquity/Ultima Global instrumentation using a 30 minute LC gradient. All MS and MS/MS data were corrected for mass drift using reference data collected from the [Glu1]-Fibrinopeptide B (human - F3261 Sigma) sampled each minute of data collection. The data were used to interrogate the E. coli BL21 database (`http://www.ebi.ac.uk/integr8`) using the full-length protein sequence using ProteinLynx Global Server v 2.5.1.

2.8.6 Protein quantification

Protein concentration in a particular sample was measured using either bicinchoninic acid (BCA) assay or measurement of absorbance of light at 280 nm.

BCA assay

This assay was particularly useful when measuring the concentration of proteins in detergent because it tolerates the presence of typical concentrations of most ionic and anionic detergents in a sample. This method works on the principle that Cu^{+2} will be reduced to Cu^{+} by proteins in an alkaline medium (the biuret reaction). The colorimetric detection of the cuprous cation is done using a reagent containing BCA. Each cuprous ion chelates with two BCA molecules and this complex exhibits a deep purple colour which can be measured at 562 nm as it has a linear absorbance with increasing concentrations. The assay was carried out using Pierce^{TM}BCA Protein Assay Kit as per the manufacturer's instructions.

Absorbance at 280 nm

A 3 mm quartz cuvette with an internal volume of 50 μL was used to measure the absorbance at 280 nm. This was done using Cary® 50 UV-Vis spectrophotometer. The extinction coefficient was computed using the ExPASy ProtParam tool (`http://web.expasy.org/protparam/`) (Wilkins et al., 1999) and the protein concentration was subsequently calculated using the Beer-Lambert law:

$$A_{280} = \varepsilon c l$$

where,

ε = extinction coefficient

A_{280} = absorbance at 280 nm

c = concentration in mol dm^{-3}

l = path length in cm.

2.9 Peptide Synthesis and Purification

2.9.1 Synthesis of Neu* transmembrane (TM) domain

Four peptides corresponding to the transmembrane domain of the oncogenic Neu* receptor with the sequence `RASWVTFIIATVEGVLLFLILVVVVGILIKRRR` were synthesised using solid-phase chemistry. The N-terminus was acetylated and the C-terminus was amidated to mimic the peptide bonds found in the parent sequence. Peptides were synthesised using solid phase 9-fluorenylmethyl carbamate (FMOC) chemistry (Carpino and Han, 1972) at the Yale University W.M KECK Facility (New Haven, CT, USA). A single amino acid per peptide was uniformly labelled with ^{13}C and ^{15}N isotopes. The position of these amino acids is shown in Table 2.9 and the rationale behind this choice is discussed in section 3.3.

Peptide name	Peptide sequence
DIX13553	acetyl-RASWVTFIIATVEG*VLLFLILVVVVGILIKRRR-conh2
DIX13554	acetyl-RASWVTFIIATVEGVLL*FLILVVVVGILIKRRR-conh2
DIX13555	acetyl-RASWVTFI*IATVEGVLLFLILVVVVGILIKRRR-conh2
DIX13556	acetyl-RASWVTFIIAT*VEGVLLFLILVVVVGILIKRRR-conh2

Table 2.9: List of Neu* peptides used in this study. Amino acids which have been ^{13}C- and ^{15}N-labelled are marked with an asterisk (*).

2.9.2 Synthesis of bovine papillomavirus (E5) transmembrane (TM) domain

The putative transmembrane domain of the oncoprotein E5 was synthesised by Insight Biotechnology UK using solid phase 9-fluorenylmethyl carbamate (FMOC) chemistry (Carpino and Han, 1972). As with the Neu* peptides, the N-terminal was acetylated and the C-terminal was amidated to mimic the peptide bonds found in the parent sequence. Additionally, non-native lysines were added on each terminal to aid solubility. The sequence of the peptide was

acetyl-KKKFLGLVAAMQLLLLLFLLLFFLVYWDHK-conh2.

2.9.3 Purification of synthetic peptides

The manufacturers supplied the peptides as crude reaction products. These could contain impurities such as deletion products from the synthesis or peptides missing one or two amino acids. Pure peptides with the molecular weight of interest, as determined by mass spectrometry (section 2.8.5), were obtained by performing reverse-phase high-performance liquid chromatography (HPLC).

Samples for injection were prepared by dissolving ∼5 mg of crude peptide in 200 μL trifluroacetic acid (TFA), 200 μL 2,2,2-trifluroethanol (TFE), 400 μL isopropanol (IPA) and then drop-wise adding a 7 : 3 mix of to IPA to a final volume of 2 mL. The sample was then run through a semi-preparative Jupiter C4 5u (300 Å, 250 × 10.0 mm) reverse-phase HPLC column (Phenomenex, Macclesfield, Cheshire, UK) connected to a purpose built two pump HPLC system (Jasco UK, Great Dunmow, Essex, UK) at a flow rate of 1.5 mL/min. A gradient constituting dH_2O and

IPA was employed as illustrated in Tables 2.10 & 2.11 for E5 and Neu* respectively. Mobile phase buffers A and B in the case of E5 are the same as those used in the purification of PDGFβR-TM (section 2.6.3). The column was equilibrated with the initial conditions beforehand.

Time (minutes)	Buffer A (%)	Buffer B (%)
0	70	30
15	70	30
50	0	100
70	0	100
90	70	30
100	70	30

Table 2.10: HPLC gradient for the purification of E5

Time (minutes)	Water (%)	Isopropanol (%)
0	70	30
15	70	30
25	60	40
50	85	15
55	100	0
60	100	0
65	70	30

Table 2.11: HPLC gradient for the purification of Neu*

The absorbance of the eluate was measured at a wavelength of 280 nm at which aromatic amino acids within the peptides absorb UV light. Multiple runs of HPLC were performed until a sufficient amount of pure protein was obtained. The presence and purity of the peptide of interest was determined using mass spectrometry (as described in section 2.8.5.

2.10 Protein Reconstitution into Detergent Micelles or Liposomes for NMR analyses

2.10.1 Reconstitution of Neu* into lipid vesicles for solid-state NMR analysis

For each sample, the fractions containing the pure peptides from HPLC were collected together separate round bottomed flasks. Their absorbances were measured and the Beer-Lambert law was used to calculate the concentration of each peptide (section 2.8.6). Volumes containing 5 mg of each of the peptides were added to a separate round bottomed flask. To these peptides, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) lipids (dissolved in 1 mL of TFE) and cholesterol (dissolved in 500 μL of chloroform) were added at a protein : cholesterol : DMPC ratio of 1 : 2 : 20.

All organic solvents were then evaporated off on a rotary evaporator while water was simultaneously added. This allowed for the creation of proteoliposomes in an aqueous environment. The absence of any residual organic solvents was confirmed using ^{1}H measurements by solution-state NMR. Approximately 5 mL of dH_2O was then added and the mixture was sonicated to break down any large aggregates.

This suspension was then transferred to a 50 mL clear centrifuge tube and freeze-thawed 10 times to facilitate the creation of unilamellar liposomes. It was then centrifuged at 4 °C and 19,000 rpm (45,000 $\times$ g) for 45 minutes. The supernatant was removed and 3 mL of 50 mM phosphate buffer (pH 7.4) was added. The mixture was centrifuged again under the same conditions as above.

The resulting pellet was resuspended in $\sim$5 mL of 50 mM phosphate buffer (pH 7.4) and transferred to 1.5 mL ultracentrifuge tubes (Beckman Coulter, High Wycombe, UK). This was followed by ultracentrifugation at 70,000 rpm (267,000 $\times$ g) at 4 °C for 15 minutes in an Optima^{TM}TLX ultracentrifuge (Beckman Coulter, UK). Any supernatant was removed with the aid of a filter paper. This step was repeated and

the sample stored in a fridge at 4 °C. Prior to the NMR experiments, this sample was funnelled into a 3.2 mm or 4 mm Magic Angle Spinning (MAS) rotor through centrifugation.

2.10.2 Reconstitution of PDGFβR-TM into lipid vesicles for solid-state NMR analysis

After purifying PDGFβR-TM (as outlined in section 2.6.2) and obtaining it in a detergent environment, the N-terminal His-tag was cleaved off using TEV protease. This was done by adding freshly purified TEV protease to the protein solution in a 1 : 5 protein to TEV protease ratio. This was left to incubate overnight at room temperature. The His-tag and the TEV protease were then removed from solution by passing it through an IMAC column. Either 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) or 1,2-dierucoyl-*sn*-glycero-3-phosphocholine (DEPC) lipids were then added to the solution such that the final protein to lipid concentration was 1 : 20. The solution was then vortexed briefly to dissolve the lipids through the formation of mixed micelles.

The detergent was removed from the solution by adding Bio-Beads^{TM}SM-2 Resin (Bio-Rad, UK). These are neutral macroporous polymeric beads and can be used to remove detergents from solution, even if they have a very low critical micelle concentration (CMC) and hence Bio-Beads are particularly useful for removal of detergents such as Triton X-100 and SDS (Rigaud et al., 1998). Bio-Beads work on the principle that they contain pores which trap detergents, however, large molecules such as proteins and lipids are not trapped and therefore detergent is selectively removed.

Prior to detergent removal, the Bio-Beads were washed with ethanol and water as per manufacturer's instructions. They were then added to the clear protein/lipid/detergent mixture and left to rotate on an end-over-end mixer overnight at 4 °C. As the detergent is removed, the turbidity of the solution increases. The detergent removal was monitored with ^{1}H measurements by solution-state NMR. The resulting sus-

pension is pelleted and proteoliposomes are prepared for solid-state NMR analysis as described in section 2.10.1.

2.10.3 Reconstitution of PDGFβR-TM into detergent micelles for solution-state NMR analysis

In order to perform solution-state NMR experiments, the protein powder obtained from the purification of PDGFβR-TM (section 2.6.3) was directly dissolved in a detergent/buffer solution. Various detergents and buffer conditions were screened (see section 6.3 for details) and optimal conditions were found at 0.4 mM uncleaved PDGFβR-TM, 100 mM n-dodecylphosphocholine (DPC), 10 mM 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC), 20 mM NaCl in 30 mM sodium phosphate buffer with pH 5.8 at 318 K.

2.11 Circular Dichroism

Circular Dichroism (CD) is a useful technique for the characterisation of secondary structure of proteins. CD measures the difference in absorbance of left- and right-handed circularly polarised light and it is this difference which is reflected in a CD spectrum. In order to estimate the secondary structure, this spectrum is compared to a database of previously obtained spectra for α-helical, β-sheet and random coil proteins (Kelly and Price, 2000; Bulheller et al., 2007).

CD experiments were carried out at room temperature ($\sim$25 °C) using a Jasco J-815 spectropolarimeter (Jasco UK, Great Dunmow, Essex, UK) and 1 mm path-length quartz cuvettes (Starna; Optiglass Ltd., Hainault, UK) requiring a sample volume of 200 μL. Spectra were recorded in the far UV region between 190 and 260 nm, with a data pitch of 0.2 nm, a 1 nm bandwidth, 50 nm/min scanning speed and a response time of 2 seconds.

CD experiments were performed on samples which contained $\sim$0.25 mg/mL of the peptide in a buffer (20 mM sodium phosphate, 20 mM NaCl, pH 7.4). The final

spectrum was obtained by averaging 32 individual spectra and subtracting a blank spectrum which is exactly the same as the sample of interest sans peptide. In certain instances where the high tension [HT] voltage exceeded 600 mV at wavelengths below 200 nm, the data were truncated to 200 nm as this indicates light scattering and potentially unreliable results. On occasion, the problem was resolved by slight lowering the concentration or using a cuvette of different path length.

2.12 Cross-Linking

The two cross-linkers used in this study were bis(sulfosuccinimidyl)suberate (BS^3) (Thermo Scientific, UK) and gluteraldehyde. All cross-linking reactions were optimised and performed as per manufacturer's instructions. This typically involved adding 20- to 50-fold molar excess of the cross-linker to the protein solution to a final concentration of cross-linker at 0.25–5 mM. The reaction mixture was incubated at room temperature for 30 minutes or on ice for 2 hours. It was then quenched with 1 M Tris-HCl (pH 7.4) such that the final concentration of Tris was 20–50 mM. The quenching reaction was allowed to continue for 15 minutes before the reaction mixture was run on gels as described in section 2.8.1.

2.13 Magic Angle Spinning Solid-State Nuclear Magnetic Resonance (NMR) Spectroscopy

Solid-state magic angle spinning (MAS) NMR experiments were conducted on 500 MHz Bruker Avance III, 600 MHz Bruker Avance II+ or 850 MHz Bruker Avance III solid-state NMR spectrometers (Bruker, Karlsruhe, Germany), where MHz indicates ^{1}H Larmor frequency. All spectrometers were equipped with 4 mm and 3.2 mm MAS probes (Bruker) capable of running in double or triple resonance modes. Samples were cooled to required temperatures using Bruker BCU Xtreme cooling unit. Spinning frequency was maintained by a Bruker MAS II controller unit. Chemical shift referencing was done with respect to the carbonyl peak of natural abundance

alanine (177.9 ppm) with respect to DSS (Morcombe and Zilm, 2003; Harris et al., 2008). Samples were prepared as detailed in sections 2.10.1 and 2.10.2. All data acquisition and initial processing was done using Topspin 2.1. All post-acquisition processing was done using Topspin 3.0 and further data analysis was done using either SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) or CcpNmr Analysis (Vranken et al., 2005) software for Mac OS X.

2.13.1 1D ^{13}C NMR experiments

Data were collected with an acquisition time of 40 ms using 4k complex data points for 64 co-added transients and a spectral window of 397 ppm. Typically, a ^{1}H–^{13}C cross polarisation (CP) with a contact time of 1 ms was used, along with 100 kHz SPINAL-64 proton decoupling during acquisition. Data was zero-filled to 65k was used along with an exponential multiplication (EM) line broadening of 100 Hz. Unless otherwise stated, spectra were acquired using a 3 s recycle delay and a $\frac{\pi}{2}$ excitation of ^{1}H and ^{13}C at pulse lengths of 2.5 μs and 4 μs respectively.

2.13.2 2D ^{13}C–^{13}C dipolar-assisted rotational recoupling (DARR) NMR experiments

All DARR experiments (Takegoshi et al., 2000) were conducted through cross-polarisation of ^{13}C nuclei using a (50–100%) ramped proton CP pulse of 1 ms with radio frequency (RF) field strengths of 90 kHz for ^{1}H and 80 kHz for ^{13}C. SPINAL-64 (Fung et al., 2000) was employed as a method of choice for the requisite proton decoupling. Experiments were conducted at either short ($\sim$30 ms) or long ($\sim$400 ms) mixing times in order to probe spin interactions over short as well as long distances in space. Data was collected using 1994 complex data points in F_2 and 336 increments in F_1. Typically, data were Fourier transformed using EM window function with line broadening of 150 Hz and a zero-filling of 1k and 4k in F_1 and F_2 respectively.

2.14 Solution-State NMR Spectroscopy

All protein NMR experiments were carried out on a 700 MHz Bruker Avance II solution-state spectrometer (Bruker, Germany) fitted with a cryoprobe (Bruker, Germany) running in triple resonance mode for $^{1}H/^{13}C/^{15}N$ experiments. All samples were analysed in 3 mm NMR tubes (Hilgenberg, Germany) and ^{1}H referencing was done with respect to water (∼4.7 ppm) on the DSS scale (Wishart et al., 1995). All samples contained 10% D_2O so that the spectrometer field was locked on to the resonance position of deuterium during acquisition. All data acquisition and initial processing was done using Topspin 2.1. All post-acquisition processing was done using Topspin 3.0 and further data analysis was done using either SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) or CcpNmr Analysis (Vranken et al., 2005) software for Mac OS X.

2.14.1 1D ^{1}H NMR experiments

All 1D ^{1}H NMR experiments were collected using the watergate method for the suppression of ^{1}H signal arising from water (Liu et al., 1998). Spectra were recorded using 32k complex data points and 32 co-added transients. A recycle delay of 1.5 seconds was used and data was typically acquired by optimising the $\frac{\pi}{2}$ pulse length for the excitation of ^{1}H nuclei whilst maintaining a power level of 5.9 dB. The acquisition time was 1.67 s and the spectral window was set to 20 ppm. Data were Fourier transformed using a zero-filling of 65k and using the exponential window function with a line broadening of 1 Hz unless otherwise specified.

2.14.2 2D $^{1}H/^{15}N$ NMR experiments

Double resonance heteronuclear single quantum coherence (HSQC) experiments (which correlate the resonances of ^{1}H and ^{15}N nuclei) (Schleucher et al., 1994) were recorded over a variable range of acquisition parameters so as to achieve the best possible spectra in the minimum time. Typically data were acquired with 2k com-

plex data points in F_2 and 128 increments in F_1 with 32 co-added transients and 16 dummy scans. A spectral width of 14 ppm was used in the ^{1}H dimension and 40 ppm was used in the ^{15}N dimension. The data were Fourier transformed using a zero-filling on 4k in F_2 and 2k in F_1. Processing was done using Gaussian multiplication (GM) window function with a line broadening of -10 Hz in both dimensions. Additionally, a polynomial baselines correction was used in F_2 and a linear prediction was used in F_1.

2.14.3 3D $^{1}H/^{15}N/^{13}C$ NMR experiments

All 3D experiments were performed as detailed in Table 2.12. All spectra were acquired at 45 °C. Processing was done by Fourier transforming using Gaussian multiplication (GM) window function and using a line broadening of -10 Hz in all dimensions. Baseline correction was done in F_1 and linear prediction was used in dimensions F_2 and F_3.

Experiment	Details	Reference
HNCO	78, 48 and 2k increments in F_1, F_2 and F_3 respectively with 36 co-added transients. Spectral widths of 22 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Kay et al., 1990)
HNCACO	92, 48 and 2k increments in F_1, F_2 and F_3 respectively with 36 co-added transients. Spectral widths of 22 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Clubb et al., 1992)
HNCA	72, 48 and 2k increments in F_1, F_2 and F_3 respectively with 40 co-added transients. Spectral widths of 32 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Kay et al., 1990)
HNCOCA	80, 52 and 2k increments in F_1, F_2 and F_3 respectively with 46 co-added transients. Spectral widths of 32 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Bax and Ikura, 1991)
CBCANH	78, 50 and 2k increments in F_1, F_2 and F_3 respectively with 56 co-added transients. Spectral widths of 75 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Wittekind and Mueller, 1993)
CBCACONH	78, 48 and 2k increments in F_1, F_2 and F_3 respectively with 32 co-added transients. Spectral widths of 75 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Grzesiek and Bax, 1992)
^{15}N HSQC–TOCSY	86, 54 and 2k increments in F_1, F_2 and F_3 respectively with 42 co-added transients. Spectral widths of 14 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Marion et al., 1989b)
^{15}N HSQC–NOESY	86, 54 and 2k increments in F_1, F_2 and F_3 respectively with 42 co-added transients. Spectral widths of 14 ppm, 40 ppm and 14 ppm in F_1, F_2 and F_3 respectively. 16 dummy scans.	(Marion et al., 1989a)

Table 2.12: Triple resonance experiments conducted in this study

2.14.4 E5 titrations

To investigate the binding of the oncoprotein E5 to the PDGFβR-TM, E5 was titrated into PDGFβR-TM reconstituted into DPC/DPPC micelles at PDGFβR-TM to E5 molar ratios of $1:1, 1:3, 1:5, 1:7, 1:9$ and $1:11$. E5 was dissolved and stored in TFE solution post-purification. The amount of E5 required per titration was dried in a 1.5 mL eppendorf tube forming a thin film. This was resuspended in PDGFβR-TM mixed micelle solution, sonicated till all the E5 dissolved and equilibrated for $\sim$1 hour. The temperature and buffer conditions were maintained and ^{1}H–^{15}N HSQC spectra recorded.

2.14.5 Deuterium exchange experiments

In these experiments amide hydrogen exchange is observed by the disappearance of ^{1}H–^{15}N resonances in an HSQC spectrum after the solvent is exchanged from H_2O to D_2O. This is done by lyophilising the sample and then resuspending it in 200 μL of D_2O followed by dissolution of the suspension through sonication. HSQC spectra are obtained before exchange and after exchange using the same experimental conditions. Additionally, a set of 5 spectra are obtained at 2 hour intervals after exchange.

Chapter 3

A Solid-State NMR Analysis of Oncogenic Neu

3.1 Introduction

As detailed in section 1.2.1, the ErbB2/HER receptor has been implicated in a very aggressive form of breast cancer which constitutes 30% of all breast cancers (Mitri et al., 2012). It has a particularly poor prognosis due to its resistance to conventional therapies and high relapse rates (Slamon et al., 1987). In order to study this oncogenic transformation, scientists have taken a keen interest in the rat analogue of this receptor known as Neu. Where ErbB2 in humans is constitutively active due to the $V_{659}E$ mutation in the TM, the oncogenic Neu mutant (hereafter referred to as Neu*) is similarly activated due to a TM $V_{664}E$ mutation. Despite being rigorously and extensively studied, there is no agreement on how a single mutation within the TM domain of the Neu* receptor affects the structural integrity of the protein.

Even though previous biophysical studies have yielded useful information on the structural effect of the point mutation in Neu*, they have largely remained inconclusive or contradictory. In this study, we aim to investigate the orientation of the Neu* TM helices with respect to each other. The insights obtained here will be particularly interesting as the investigative technique of choice, solid-state NMR spectroscopy, allows us to look at these dimers in 'native-like' lipid bilayers. Furthermore, the addition of cholesterol grants the bilayer a fluidity and heterogeneity

which make the results obtained in this section all the more biologically relevant. This is contrary to solution-state NMR where protein solubilisation is achieved by embedding the proteins in detergent micelles. This can lead to distortion of membrane protein structure due to the curvature stress induced by the micelles.

Further to addressing the specific problem described in the following section, the novel protocols that have been developed as part of this study can be used as the basis of investigating structural properties of a host of membrane proteins.

3.1.1 Previous work

Preliminary investigations of the dimeric structure of Neu* have been conducted in the Dixon/Brown groups using solid-state NMR previously (Ellis et al., unpublished data). However, the experimental procedure adopted here is different in several ways, of which two are significant:

- At no point during the sample preparation was the peptide freeze-dried so as to reduce the chances of compromising the structural integrity which may result from this process (Roy and Gupta, 2004). This has been explained in section 3.5.

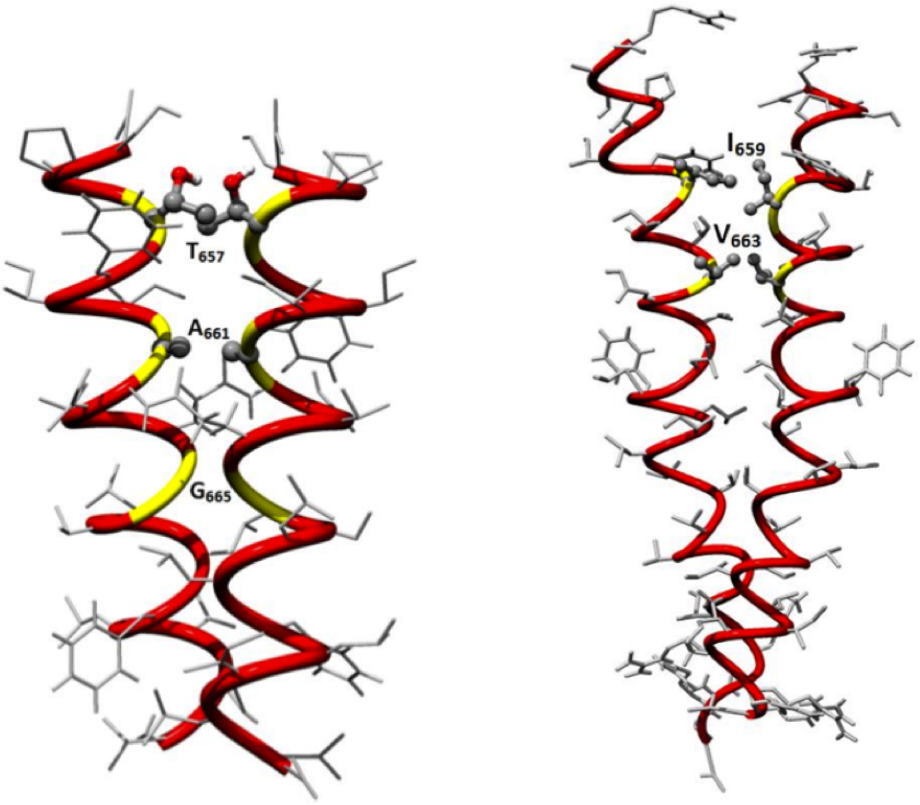
- Simultaneous studies in the Dixon group suggested that the cholesterol concentration in 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) bilayers affects the the structural properties of the membrane as a function of temperature. Importantly, it was found that at subzero temperatures, increasing the cholesterol concentration reduced bilayer rigidity. This is because addition of cholesterol results in bilayers in liquid-crystalline (L_α) phase rather than gel-phase (L_β). This was believed to be the cause for the reduction in signal to noise ratio for GpA peptide signals inserted in a DMPC/cholesterol bilayer (Patel et al., unpublished data). The cholesterol concentration in this study was hence reduced from a lipid:cholesterol ratio of $3:1$ to $10:1$, which represents a reduction of about 64%.

3.2 Models for the Neu* Dimer

The inter-helical interactions that have been hypothesised to stabilise the Neu* dimer shall be discussed here. There are two sequence motifs that are thought to pack at the interface of the helical Neu* TM domain dimer.

The first model (Model A) suggests that the $A_{661}XXXG_{665}$ motif in the transmembrane domain stabilises the dimer, with T_{657} also lying at the interface. This model has been supported by 1D REDOR ssNMR experiments and computational modelling (Smith et al., 2002b) and is illustrated in Figure 3.1a.

The second model (Model B), on the other hand, suggests that the $I_{659}XXXV_{663}$ motif stabilises the dimer with I_{659} and V_{663} lying at the dimeric interface. Studies conducted to support this model include Molecular Dynamics (MD) simulations, site-specific infrared dichroism (Beevers and Kukol, 2006), FTIR spectroscopy and mutagenesis studies (Beevers et al., 2010). This model has been structurally illustrated in Figure 3.1b.



(a) Model A: A $A_{661}XXXG_{665}$ motif stabilises the dimer

(b) Model B: A $I_{659}XXXV_{663}$ motif stabilises the dimer

Figure 3.1: Models for Neu* TM dimer packing

3.3 Labelling Scheme

3.3.1 Motivation and scientific approach

To verify which of the models outlined in section 3.2 can be corroborated using NMR correlation experiments, pairs of synthetic peptides were designed to be used together to test each model. Synthetic peptides are particularly useful in NMR analyses as their use allows amino acids at very specific positions to be labelled. Peptides can be mixed together and due to one amino acid per peptide chain, interactions can be confirmed to be inter-molecular as opposed to intra-molecular. In this case the amino acids chosen are those which are close to the proposed dimer interface.

In designing each pair of peptides, the following considerations and/or requirements

were important:

- For each of the helices within a particular model, an amino acid at or very near the suggested dimeric interface would be uniformly ^{15}N- and ^{13}C-labelled.
- Short inter-helical distances of up to 5Å between the pair of labelled amino acids must be chosen to detect through-space intermolecular magnetisation transfer.
- Average chemical shifts of each of the resonances of the selected amino acids must be as different as possible to its counterpart in the other peptide. This is to ensure that there are no overlapping peaks in the NMR spectrum. This becomes all the more important when conducting the experiments at low temperatures where the line-broadening observed due to conformational inhomogeneity is introduced by quenching of motions upon freezing.

3.3.2 Design

With the aforementioned considerations in mind, PDB files of each of the models were used to guide the design of synthetic peptides for solid-state NMR analyses. This approach avoids the spectral crowding that can occur due to uniformly labelled samples, especially when experiments are conducted at low temperatures and peaks tend to be broad. As mentioned earlier, to verify each model, a pair of peptides will be used together to check for inter-helical magnetisation transfer.

Model A

The first pair of peptides was designed to detect inter-helical magnetisation transfer across the $A_{661}XXXG_{665}$ dimeric interface as suggested by Model A. Many short inter-helical distances between carbon atoms were predicted by this model (see Table 3.1) including a distance of 3.90 Å between Cα of G_{665} and Cδ_2 of L_{668} and 3.59 Å between CO of G_{665} and Cδ_2 of L_{668} (Figure 3.2a). In light of this, two peptides were designed: one with [U–^{15}N, ^{13}C]-G_{665} as the only isotopically labelled

amino acid in the sequence and, similarly, the other with [U–^{15}N, ^{13}C]-L$_{668}$ as the only isotopically labelled amino acid in the sequence.

Model B

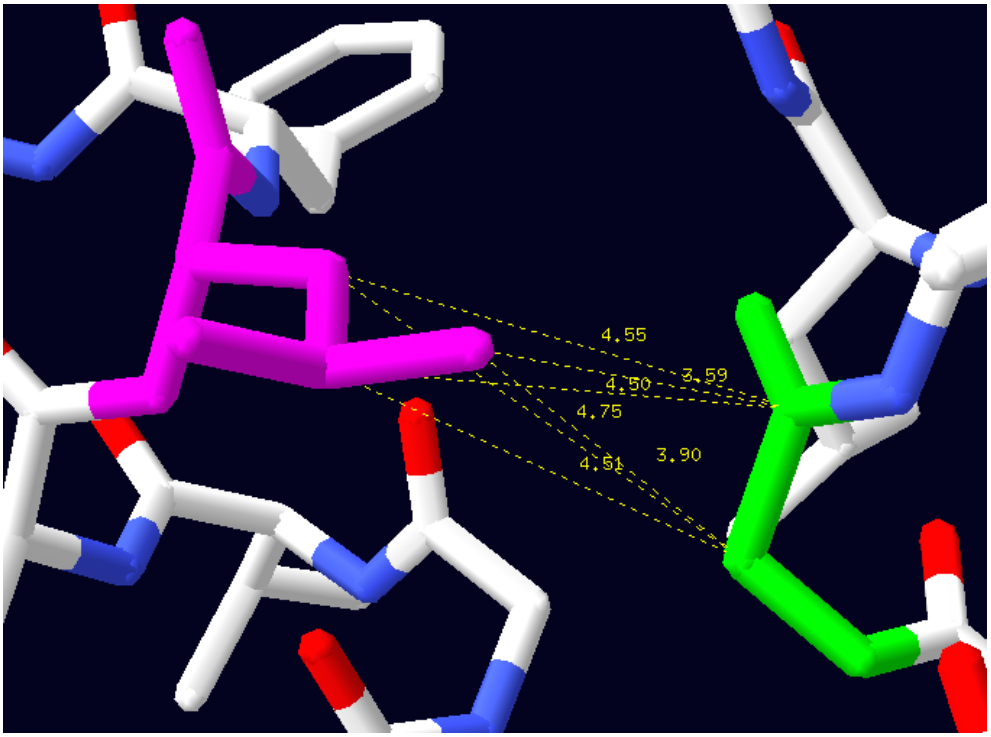
Similarly, short inter-helical distances of 4.35 Å between the Cγ_2 of I$_{659}$ and the Cβ of T$_{662}$ were predicted in Model B. Additionally, 4.18 Å between the Cγ_2 of I$_{659}$ and the Cγ of T$_{662}$ was predicted (see Table 3.2 and Figure 3.2b). In this case, however, an additional motivation behind choosing I$_{659}$ was the FTIR spectroscopy and mutagenesis data reported previously (Beevers et al., 2010) which suggested a strong role of I$_{659}$ in oligomerisation.

Table 3.1: List of predicted inter-atomic distance between L$_{668}$ and G$_{665}$. Distances less than 5 Å are highlighted in green.

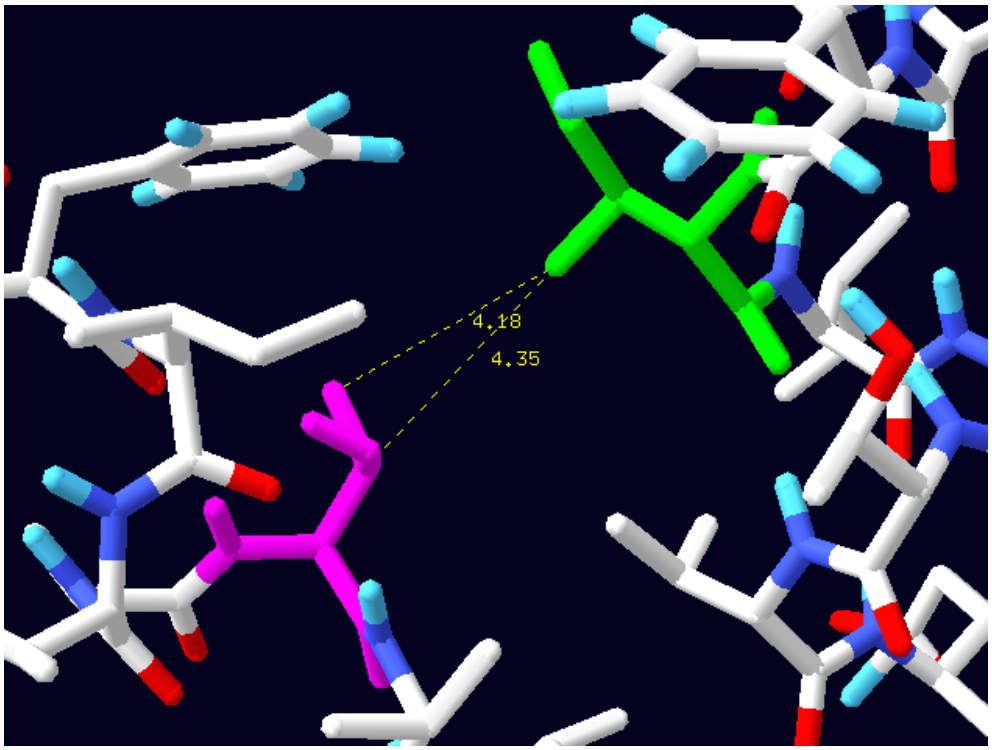
Leu/Gly	CO	Cα
CO	6.42	6.65
Cα	5.93	5.97
Cβ	4.55	4.75
Cγ	4.50	4.51
Cδ_1	5.91	5.95
Cδ_2	3.59	3.90

Table 3.2: List of predicted inter-atomic distance between I$_{659}$ and T$_{662}$. Distances less than 5 Å are highlighted in green.

Ile/Thr	CO	Cα	Cβ	Cγ
CO	7.75	7.75	6.45	6.81
Cα	7.95	7.74	6.30	6.50
Cβ	7.83	7.33	5.83	5.67
Cγ_1	8.52	7.36	6.33	6.00
Cγ_2	6.42	5.85	4.35	4.18
Cδ	9.38	8.54	7.02	6.38

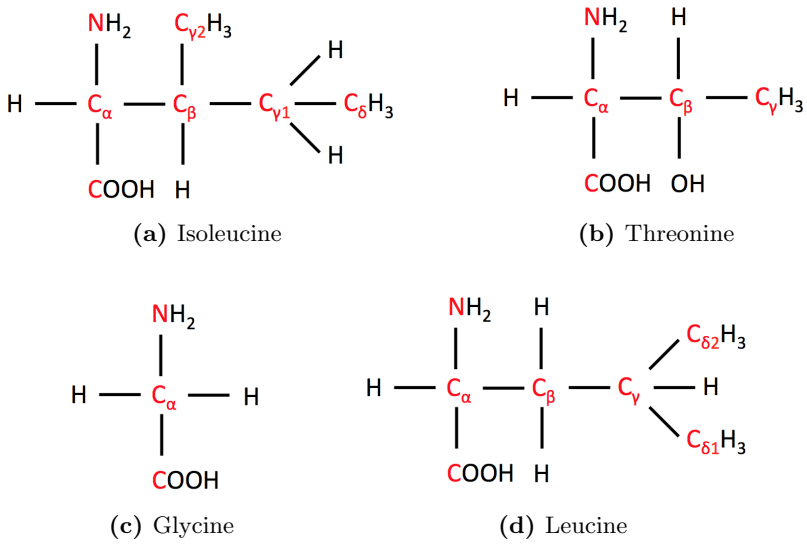


(a) Model A: Distances less than 5 Å are shown between L_{668} (magenta) and G_{665} (green).



(b) Model B: Distances less than 5 Å are shown between T_{662} (magenta) and I_{659} (green).

Figure 3.2: Predicted inter-atomic distances for Neu* TM dimer packing for each of the models.



(a) Isoleucine **(b)** Threonine

(c) Glycine **(d)** Leucine

Figure 3.3: Chemical structures of uniformly ^{15}N- and ^{13}C-labelled amino acids. All labelled nuclei are in red.

The labelling scheme hence adopted is represented in the following peptide sequences, where (*) represents a labelled residue:

Model A

Peptide pair to test: acetyl-RASWVTFIIATVEG*VLLFLILVVVVGILIKRRR-conh2

acetyl-RASWVTFIIATVEGVLL*FLILVVVVGILIKRRR-conh2

Model B

Peptide pair to test: acetyl-RASWVTFI*IATVEGVLLFLILVVVVGILIKRRR-conh2

acetyl-RASWVTFIIAT*VEGVLLFLILVVVVGILIKRRR-conh2

The labelled amino acids are shown in Figure 3.3, with isotopically labelled nuclei shown in red. After the peptides were purified, the pair for each model was combined (to form dimers) and reconstituted into a lipid bilayer for NMR analyses.

3.4 Peptide Purification

Peptides were provided by the manufacturer in a crude form containing truncated products and protecting groups from the solid-phase synthesis. These were purified

using reverse-phase HPLC (RP-HPLC) as detailed in section 2.9.3 and then analysed by mass spectrometry.

3.4.1 High performance liquid chromatography (HPLC)

The crude protein was dissolved and passed through a column containing a porous solid matrix. Depending on the strength of the interactions that the proteins (mobile phase) have with this solid matrix (stationary phase), they are retained and the rate at which they pass through the column will vary accordingly, hence separating different peptides. A characterising feature of an HPLC column is the length of the hydrophobic alkyl chains (-CH_2-CH_2-CH_2-CH_3) which form the stationary phase and that interact with the analyte. Commonly used chain lengths are C4, C8, and C18. C4 is generally used for proteins and C18 is generally used to capture peptides or small molecules. The eluted solution then passes through a UV-Visible absorption detector before the fractions are collected. Peptides absorb at wavelengths of 220 nm due to C = O bonds, and if amino acids with aromatic side chains, such as tryptophan, are present, they also absorb more strongly at wavelengths of 280 nm due to C = C bonds. This produces a UV-Vis absorption spectrum which can be used to help identify the fractions containing the desired peptide. The fractions which correspond to peaks on the chromatogram can then be characterised by mass spectrometry.

Even though a C4 column was used in this study, which is ideally suited for the purification of these extremely hydrophobic peptides, they can bind very strongly to the stationary phase of the column and may prove difficult to elute. Resultantly, clear baseline separation is lost and the peaks are broad, which often means unresolved peaks overlap and purification is not possible. This phenomenon can be controlled by optimising two parameters during peptide purification. Firstly, the gradient of the organic solvent which is used as the mobile phase can be optimised. Secondly, the flow rate can be carefully optimised to achieve maximum resolution.

Several runs of purification of Neu* crude peptide were performed on the HPLC

instrument and the UV absorbance at 280 nm was recorded. A representative chromatogram with optimised gradient is shown in Figure 3.4. It was confirmed using mass spectrometry that the peptide eluted at about 48 minutes at a mobile phase composition of about 80% isopropanol : 20% water. The peak observed at 15 minutes is most likely due to by-products from the synthesis of the peptide. Most of these compounds are more soluble than the peptide and hence elute much earlier. The other peaks observed which are closer to the Neu* peak are likely due to deletion products from the synthesis, for example peptides missing one or two amino acids. After the peak of interest was identified, only fractions containing that peak were collected. Approximately 10 runs of about 5 mg of crude peptide were required to obtain enough peptide to fill a 4 mm MAS ssNMR rotor.

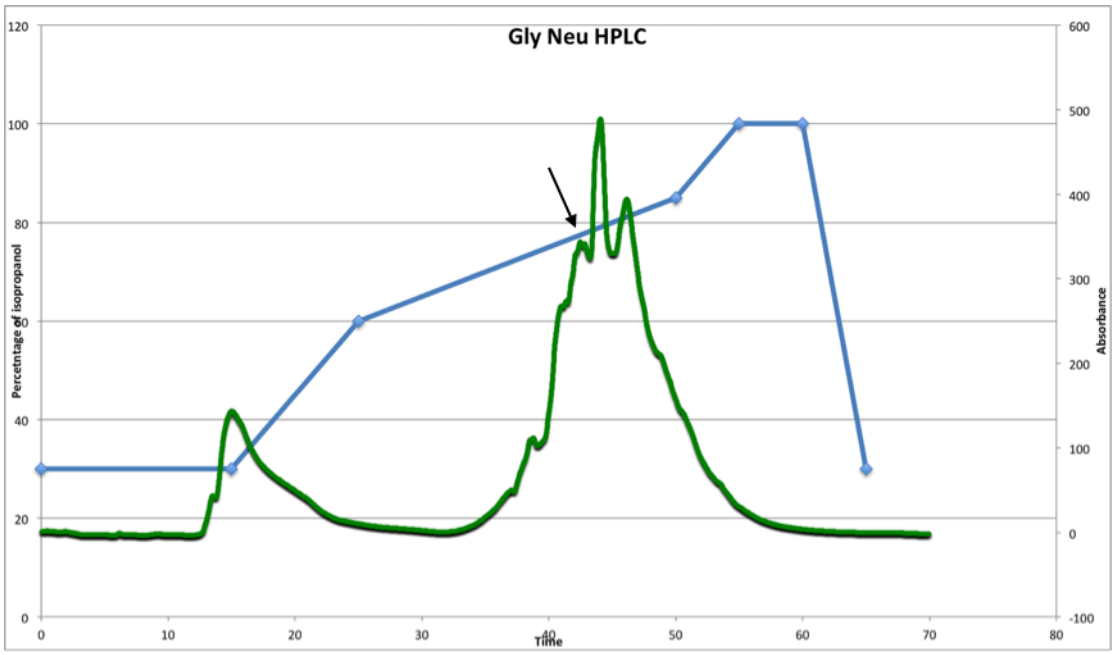


Figure 3.4: Neu* peptide was purified by a Jasco HPLC instrument using a Phenomenex Jupiter C4 column. The chromatogram shows the UV absorbance at 280 nm (green line) as the percentage of organic (isopropanol) passing through the column is varied. This gradient is illustrated in blue. The peak at about 48 minutes corresponds to pure Neu* peptide (as indicated).

3.4.2 Electrospray ionisation mass spectrometry (ESI-MS)

Mass spectrometry is an analytical technique that can be used to measure the mass-to-charge ratio of charged entities within a sample. It can be used to pro-

vide information about various physical and chemical properties of a compound. These properties include masses of particles, the elemental composition of a sample or molecule, and the chemical structures of molecules, such as peptides and other chemical compounds.

In this study, mass spectrometry was used after HPLC to identify fractions which contained the desired peptide (of which the mass is known) and determine the purity. All such characterisation was carried out on a micrOTOF instrument (Bruker, UK). This uses an electrospray ionisation source (ESI) which allows small quantities of the peptide in solution to be taken straight from the HPLC fractions. The sample passes through a nebuliser, is dried to remove solvent and then ionised.

The mass spectrum of the fraction obtained at 48 minutes for the peptide labelled at G_{665} is shown in Figure 3.5. In the mass spectrum obtained, charge states corresponding to +5, +4, +3, +2 and +1 can be observed (top panel). The total mass of the major species obtained from the deconvoluted spectrum is 3794 Da (bottom panel), which agrees well with the theoretical mass of 3795 Da for the peptide. A peak at about 3906 Da can also be observed. This is due to TFA-adducts formed during the purification process and have been observed previously (Hua and Kobertz, 2013).

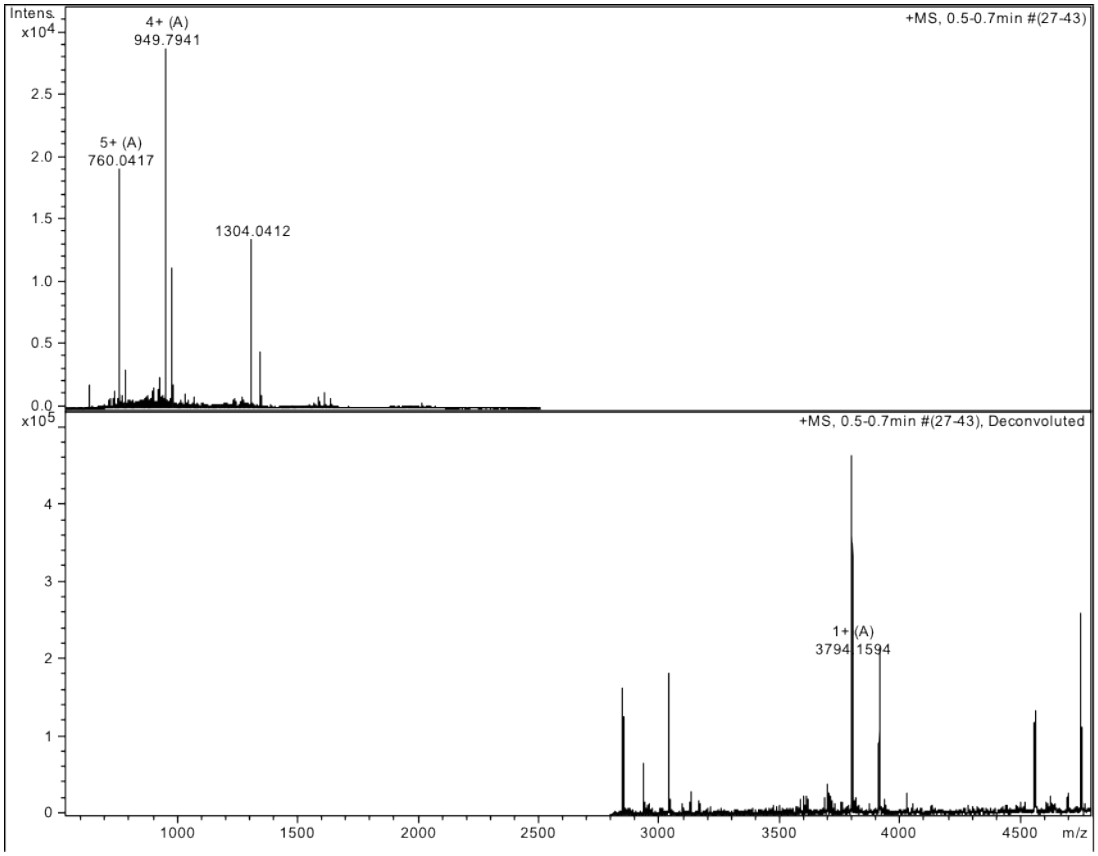


Figure 3.5: Assessment of purified Neu* through ESI-MS. Pooled fractions from the HPLC purification were checked for the existence of the peptide and it's purity. The top panel shows the mass spectrum with multiple charge states of the peptide. The bottom panel shows a major peak corresponding to the Neu* peptide (3794 Da) with an additional peak at about 3906 Da which can be attributed to a TFA adduct.

3.5 Reconstitution of Neu* peptides in lipid bilayers

Reconstitution of the Neu* peptide into DMPC/cholesterol liposomes was carried out by dissolving all three components into organic solvents, combining them together to thoroughly mix, and gradually adding phosphate buffer (pH 6.8) as organic solvents are removed using a rotary evaporator (see section 2.10.1 for details). This protocol was adopted in lieu of removing all organic solvents under vacuum to form a film that is subsequently hydrated so as to prevent the peptide from undergoing changes which may compromise its structural integrity. Specifically, previous studies have shown that structural properties such as α-helical content of a protein can be artificially altered through freeze-drying (Roy and Gupta, 2004). Even though the

molecular basis for this isn't completely understood, several reasons have been put forth. As an example, water molecules bind to a protein in a particular order: the first to interact are charged and polar amino acids, next, around the hydrophobic clusters. The former stage is particularly important because in the absence of water molecules, side chains may interact with each other and 'lock-up' certain conformations which may prevent the flexibility required for dimer formation (Poole and Finney, 1983; Falconi et al., 2003).

In order to ensure that all organic solvents were removed from proteoliposomes, ^{1}H spectra were collected to look for characteristic ^{1}H chemical shifts of 2,2,2-trifluoroethanol (TFE) and isopropanol. Figure 3.6 shows a representative spectrum in which no solvent peaks are observed. The peaks that can be observed upfield of 4 ppm and the single peak at 8.17 ppm can be attributed to cholesterol and lipids in the solution. The negative peak at 4.711 ppm is due to WATERGATE water suppression.

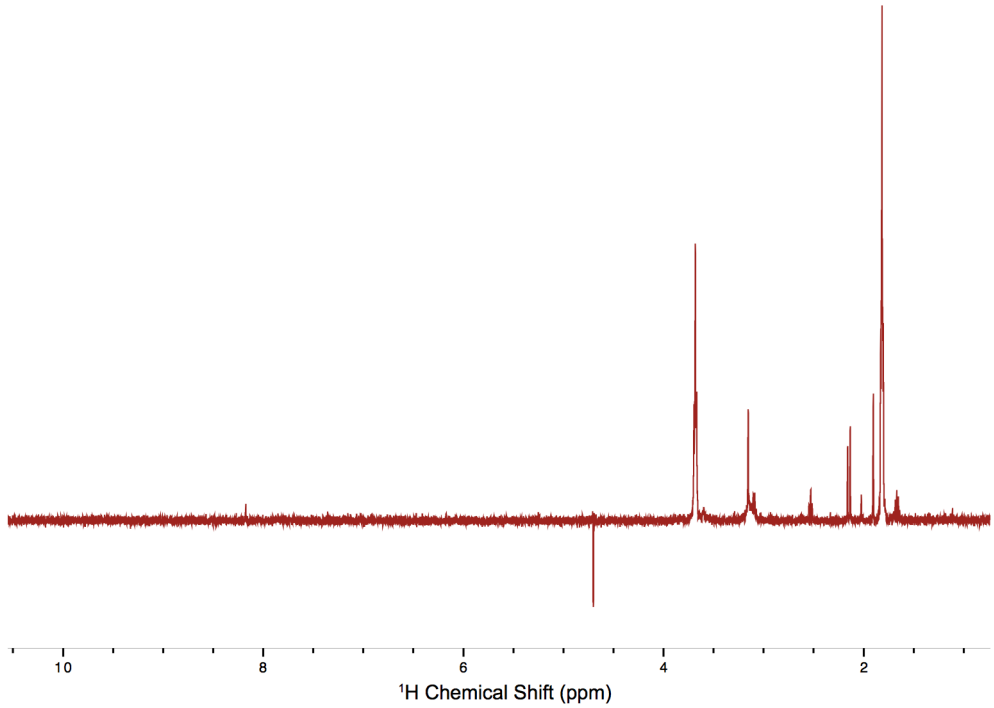


Figure 3.6: ^{1}H NMR spectrum of Neu* proteoliposomes. No peaks which can be attributed to organic solvents are observed.

3.6 Circular Dichroism (CD)

To evaluate the secondary structure of the Neu* peptide in DMPC/cholesterol bilayers, CD spectra were obtained as per the protocol detailed in section 2.11. CD measures the difference in absorbance of left- and right-handed circularly polarised light and it is this difference which is reflected in a CD spectrum.

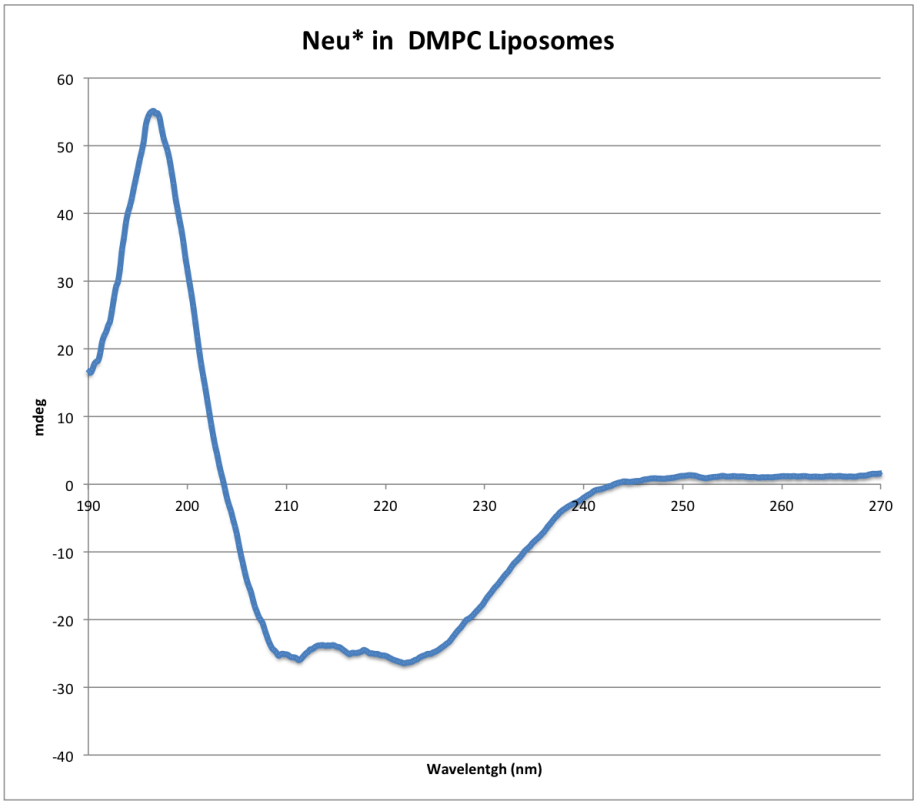


Figure 3.7: CD spectrum of Neu* in DMPC liposomes recorded between 190–270 nm at a peptide concentration of 0.1 mg/mL. The characteristic negative peaks at around 208 nm and 222 nm are indicative of α-helical structure of Neu* in DMPC liposomes.

Samples for analysis were prepared as described in section 2.10.1 and diluted with 50 mM phosphate buffer (pH 7.4) to a final protein concentration of ~0.1 mg/mL. The protein : cholesterol : DMPC molar ratio was maintained at 1 : 5 : 15. An examination of the CD spectrum reveals two minima at 210 nm and 223 nm and a maximum at about 197 nm, all of which are characteristic of an α-helical structure in membranes (Wallace et al., 2003). These are slightly shifted from values observed in soluble proteins (nminima at 208 nm and 222 nm whereas a maximum is seen at

192 nm). These shifts have been observed previously in the Dixon group and in the scientific literature (Wallace et al., 2003) for a host of membrane peptides. Several explanations for this phenomenon have been put forth, including the hypothesis that detergent micelles or bilayer environments posses a low dielectric constant which can affect the relative ground and excited states of the electronic transitions of the backbone of membrane embedded peptides relative to those which are soluble (Cascio and Wallace, 1995; Chen and Wallace, 1997).

3.7 Solid-state NMR

Samples for analysis were prepared as described in section 2.10.1. Rotors were placed in appropriate probes for magic angle spinning (MAS).

3.7.1 Probe CP setup

Prior to each measurement of the Neu* samples, a ^{13}C cross-polarisation (CP) spectrum was obtained on natural abundance alanine for referencing. Additionally, the COOH, CH and CH_3 peaks provided magic angle, shimming and decoupling controls respectively. It was ensured that the COOH, CH and CH_3 linewidths were less than 30 Hz, 50 Hz and 30 Hz respectively. In order to monitor the health and status of the probe, the parameters for acquisition of alanine CP were kept constant at 16 scans, 10 kHz MAS, 40 ms acquisition time, 1 ms contact time, 100 kHz TPPM decoupling, 70–100% ramp for CP (on ^{1}H) and 3 s recycle delay. Figure 3.8 shows a ^{13}C alanine cross-polarisation spectrum obtained on a 500 MHz spectrometer in a 4 mm rotor.

3.7.2 Temperature

A low-temperature of -15 °C was chosen to conduct all MAS ssNMR experiments. At this temperature, the DMPC lipids are in 'gel' phase (Needham and Evans, 1988). The rationale behind the low temperature was that this would lead to reduction of

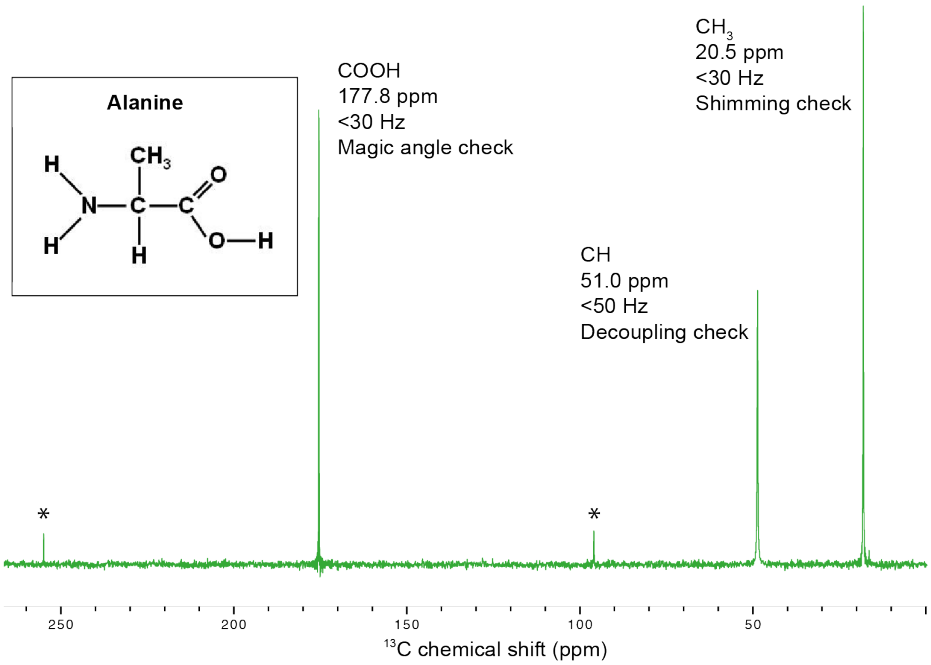


Figure 3.8: ^{13}C alanine cross-polarisation spectrum. The inset shows the chemical structure of alanine. Each of the carbon peaks are labelled. Spinning sidebands are labelled with an asterisk (*).

internal motions within the protein (such as side-chain dynamics), which in turn would improve spectral resolution (Abdine et al., 2010). However, this would also lead to line-broadening as dipolar couplings and T_1 relaxation times are increased (Hiller et al., 2005; Frericks et al., 2006; Cady et al., 2009). Since only a single amino acid was labelled per peptide, spectral crowding due to broadening wasn't a major concern and the benefits of performing experiments at low temperatures were capitalised upon. Furthermore, cholesterol can be used to alter the rigidity of the membrane, and increasing cholesterol concentrations may allow for the experiments to be conducted at higher temperatures which are more physiologically relevant (Luo et al., 2009).

Before conducting 2D DARR experiments (explained in section 1.6.4), 1D ^{1}H spectra were obtained to corroborate that the samples were indeed frozen. Figure 3.9 shows a comparison of the peptide ^{1}H spectrum at two temperatures. The spectrum at 25 °C (pink) shows relatively sharp peaks, but as the temperature is dropped to -15 °C, almost all the peaks broaden considerably (green).

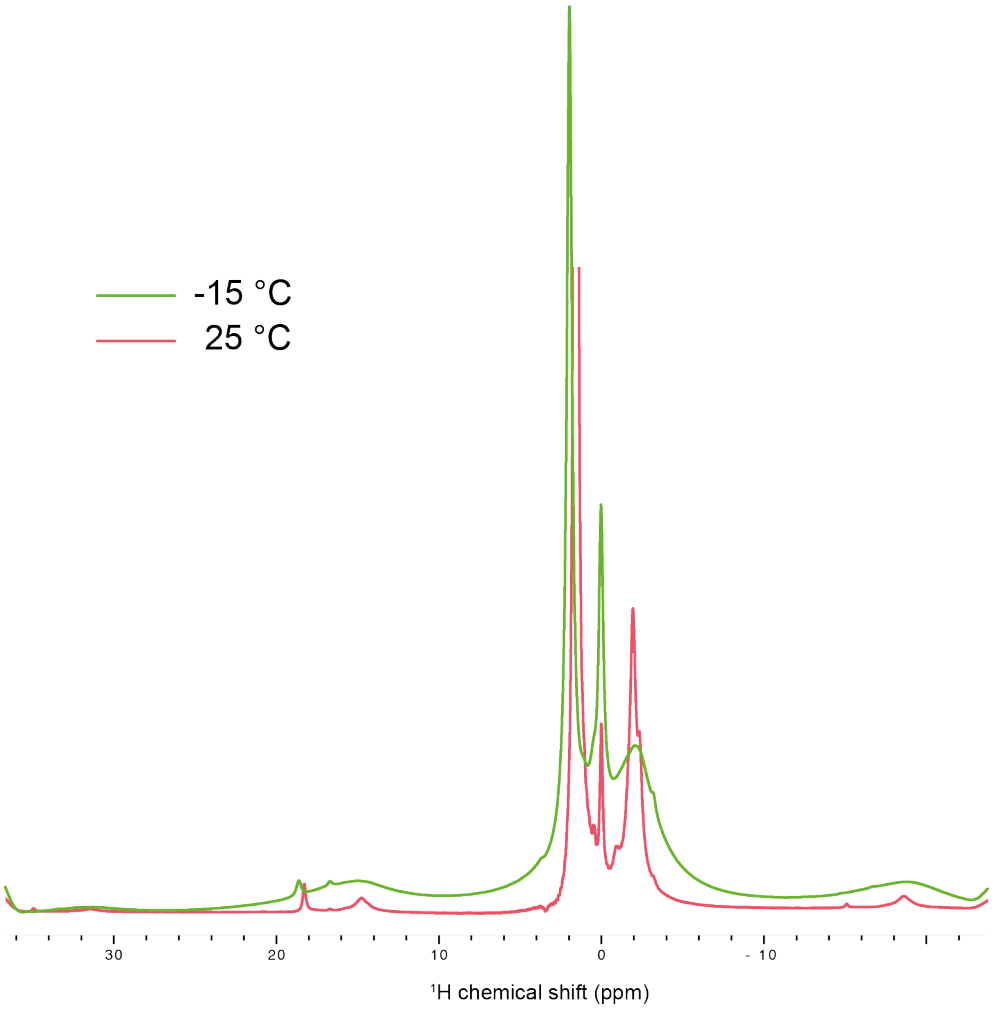


Figure 3.9: Comparison of the ^{1}H spectra of the sample made to investigate model A. The spectrum at 25 °C (pink) shows relatively sharp peaks but as the temperature is dropped to -15 °C, almost all the peaks broaden considerably (green).

3.7.3 Cross polarisation (CP)

NMR as a technique suffers from inherently low sensitivity. This is particularly true for biologically important nuclei such as ^{13}C and ^{15}N which have a low gyromagnetic ratio and natural abundance. For this reason a basic pulse-acquire solid-state NMR experiment, known as the Bloch decay (BD) experiment, yields only a small fraction of the expected signals (Kolodziejski and Klinowski, 2002). Another contributing factor to low sensitivity is the fact that spin-lattice relaxation times of dilute spin-$\frac{1}{2}$ nuclei tend to be quite long. This means long experimental times because sufficient

delays must be left between consecutive scans. This can be particularly problematic when several thousand scans are required to get a reasonable signal to noise ratio, making the acquisition of quality data impractical.

When combined with MAS, spin polarisation from a nucleus with high gyromagnetic ratio such as ^{1}H can be transferred to nuclei like ^{13}C whose gyromagnetic ratio is fourfold lower. Cross polarisation (CP) is a frequently exploited technique to achieve this transfer and relies on Hartmann-Hahn matched radio frequency (RF) pulses (see section 1.6.3) for details).

Optimisation of CP contact time

For each of the samples, the CP transfer was optimised for the best efficiency. A crucial parameter is the CP contact time describing the time during which magnetisation is transferred from the nucleus with a high gyromagnetic ratio (usually ^{1}H) to the dilute nucleus (e.g. ^{13}C). The build up of magnetisation for various ^{13}C nuclei in the biological molecules analysed in this study depends on the extent of the dipolar coupling to the proton network. The extent of this dipolar coupling depends on the degree of protonation for each type of carbon and any molecular motions (such as methyl group rotation) which may contribute to the averaging out of this dipolar coupling (Kolodziejski and Klinowski, 2002). Protonated carbons, with reduced mobility, have a faster rate of magnetisation build-up while methyl groups with higher rotational mobility build-up at a slower rate. As a result, different nuclei will cross polarise at different rates. The length of the contact time is chosen such that all types of carbons of interest have had sufficient time to polarise but not so long so as to loose significant magnetisation. In order to optimise signal to noise, a range of 1D ^{13}C CP MAS spectra of Neu* were recorded at a range of contact times. Figure 3.10 shows a range of contact times tested, a contact time of 0.8 ms is chosen as after that no improvement in polarisation transfer is observed.

Figures 3.11 & 3.12 show ^{13}C CP spectra of Neu* samples which were made to test model A and model B respectively. Assignment of all amino acids was performed

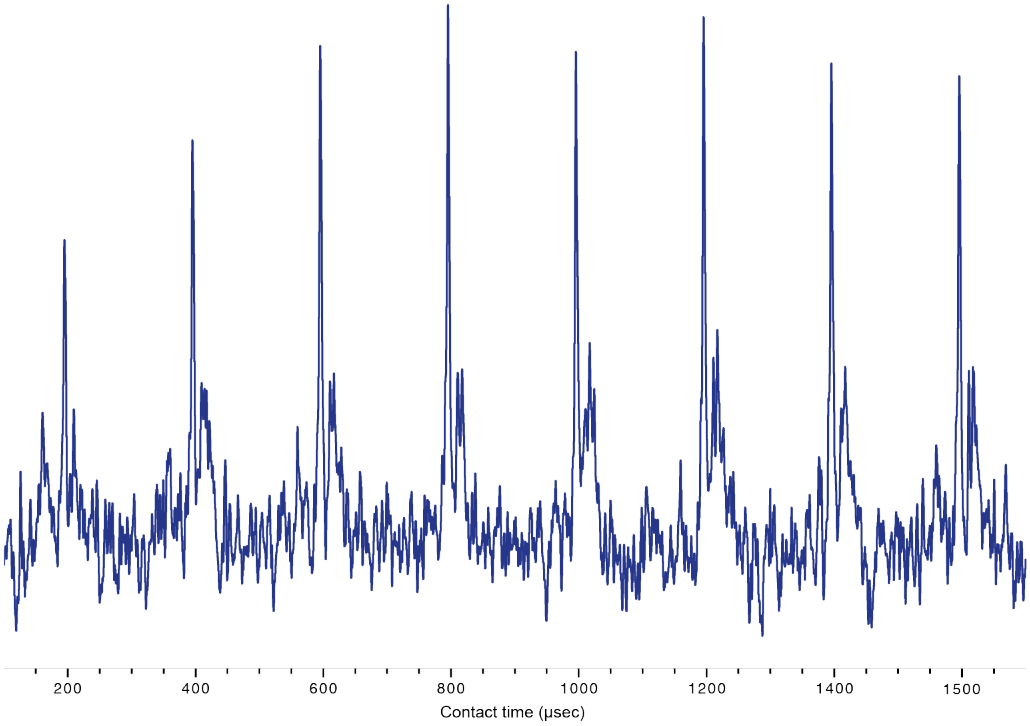


Figure 3.10: Optimisation of contact time for CP. For this particular sample, run at 850 MHz spectrometer on a U-^{13}C G_{665}/L_{668} sample. Magnetisation build-up and polarisation transfer is improved up until a contact time of 0.8 ms after which it plateaus out.

as described in section 3.8.2. Several additional peaks were also observed which were attributed to DMPC lipids present at natural abundance. These are identified using data previously acquired for DMPC liposomes in the absence (Lee and Griffin, 1989) and presence (Patel et al., unpublished data) of 5% cholesterol. The DMPC resonances typically appear as sharp peaks (0.5-1.6 ppm) due to the fast internal Brownian motions of the DMPC molecules (Saffman and Delbrück, 1975).

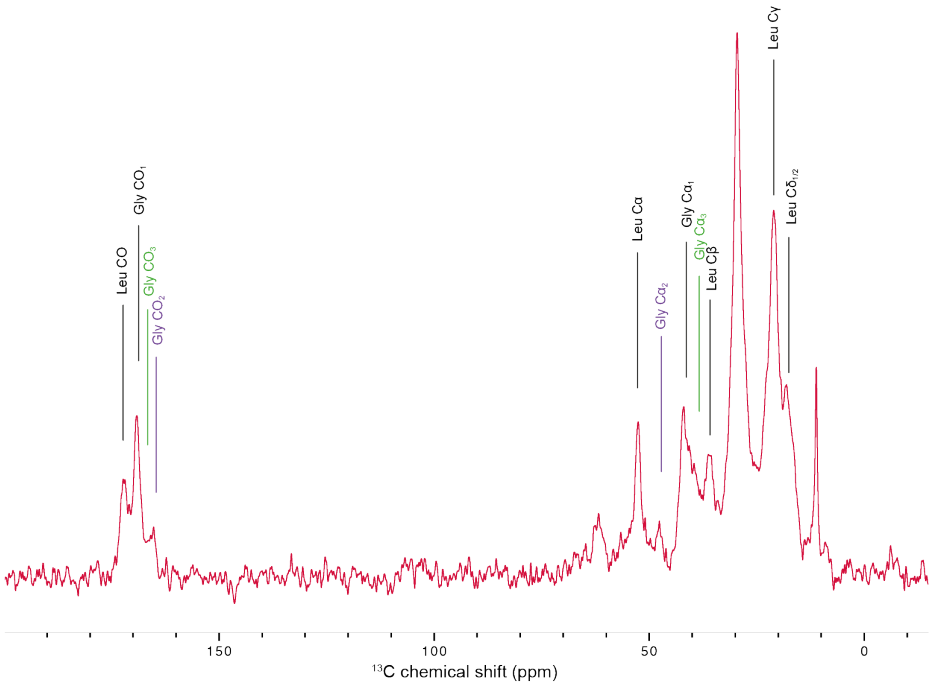


Figure 3.11: ^{13}C CP spectrum of Neu* [U–^{15}N, ^{13}C]-G_{665}/L_{668} labelled sample made to investigate model A in DMPC liposomes. Peaks which can be attributed to [U–^{15}N, ^{13}C]-G_{665} or [U–^{15}N, ^{13}C]-L_{668} have been indicated. Secondary species from G_{665} are marked in purple and tertiary species are marked in green. Spectrum recorded on a 600 MHz spectrometer with 10 kHz MAS at -15 °C for 256 co-added transients.

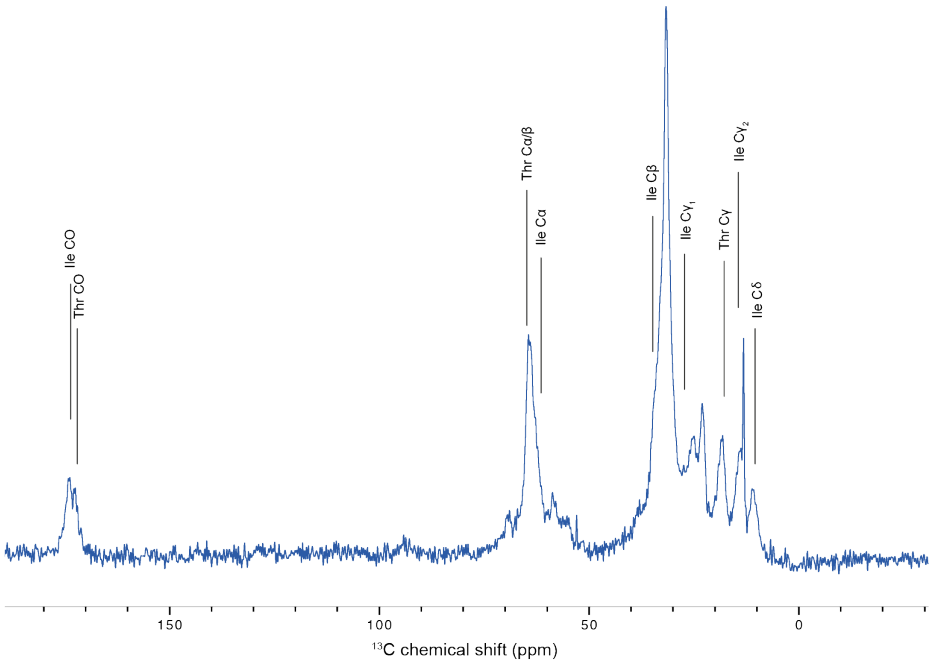


Figure 3.12: ^{13}C CP spectrum of [U–^{15}N, ^{13}C]-T_{662}/I_{659} Neu* labelled sample made to investigate model B in DMPC liposomes. Peaks which can be attributed to [U–^{15}N, ^{13}C]-T_{662} or [U–^{15}N, ^{13}C]-I_{659} have been indicated. Spectrum recorded at 500 MHz with 10 kHz MAS at -15 °C for 10,000 co-added transients.

3.7.4 ^{13}C polarisation enhancement using nuclear Overhauser polarisation (NOP)

Even though CP has been the technique of choice for polarisation enhancement for most solid-state NMR experiments for a long time, alternative techniques have been introduced more recently which offer an advantage over CP for certain applications. One such technique called nuclear Overhauser polarisation (NOP) relies on the nuclear Overhauser effect (NOE) for polarisation transfer (Takegoshi and Terao, 2002). NOE enhancement occurs for ^{13}C spins in fast-rotating groups such as CH_3. The enhanced polarisation of mobile $^{13}CH_3$ spins is then successively transferred by DARR to the other ^{13}C spins, which in turn leads to NOE enhancement of the other stationary ^{13}C spins. It has been shown that NOP transfer works particularly well for mobile biological molecules (Katoh et al., 2004).

NOP provides three main advantages over CP. Firstly, it works well with mobile molecules while CP signal enhancement becomes smaller with short $T_{1\rho}$ of ^{1}H or ^{13}C. Secondly, the repetition time of acquisition is reduced even when ^{1}H T_1 is long. Thirdly, NOP enables a more quantitative comparison of signal intensities. For these reasons, and the fact that the peptide is expected to demonstrate at least a certain degree of mobility, NOP polarisation transfer was also tried on the U-^{13}C G_{665}/L_{668} Neu* sample using the same experimental conditions as previously.

A comparison of ^{13}C NOP-MAS and CP-MAS spectra for the [U–^{15}N, ^{13}C]-^{13}C-G_{665}/L_{668} Neu* sample is shown in Figure 3.13. Even though several sharp peaks are observed for NOP, including that of L_{668} Cα, several peaks such as G_{665} Cα_1 and G_{665} CO have completely disappeared. This could be because at low temperatures the narrowing condition for methyl rotation is not met (Katoh et al., 2004). There is also a broad peak centred around the MAS sideband at about 105 ppm. These features (or lack thereof) would prove greatly detrimental to the analysis here and hence it was decided not take this approach forward.

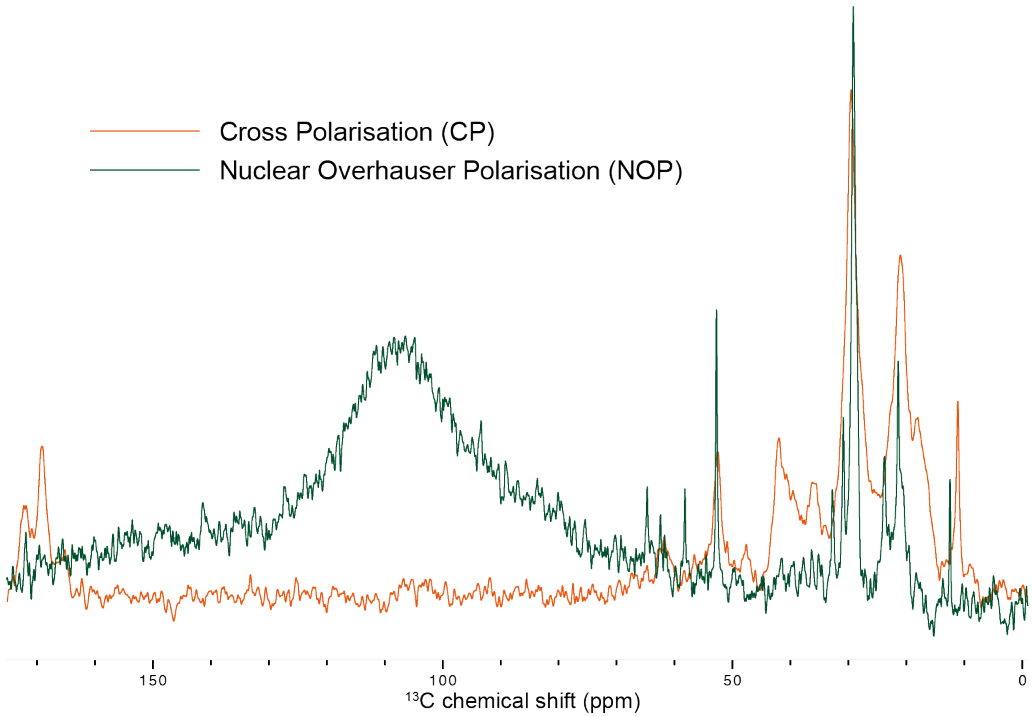


Figure 3.13: A comparison of ^{13}C NOP-MAS and CP-MAS spectra for [U–^{15}N, ^{13}C]-G_{665}/L_{668} Neu* sample. ^{13}C NOP-MAS spectrum is in green and ^{13}C CP-MAS spectrum is in orange. Spectra recorded at 850 MHz spectrometer with 10 kHz MAS at -15 °C.

3.8 Homonuclear 2D Solid-State NMR of Neu*

3.8.1 Pulse sequence

In order to investigate the dimeric structure of Neu* using labelled samples as described above, it was imperative that an RF pulse sequence which can provide the most efficient ^{13}C–^{13}C long-range through-space correlations be selected. Preliminarily, the following pulse sequences were considered:

Proton-driven spin diffusion (PDSD)

In this pulse sequence, CP is carried out followed by a $\frac{\pi}{2}$-pulse which then restores the carbon magnetisation along the z-axis. Spin diffusion is then allowed to occur between carbon atoms for a set amount of time known as the mixing time (t_{mix}). Finally, another $\frac{\pi}{2}$-pulse brings the carbon magnetisation back to the transverse plane

and the free induction decay (FID) is recorded under high power proton decoupling (Szeverenyi et al., 1982; Pines et al., 1973).

Dipolar-assisted rotational resonance (DARR)

This is a modification of PDSD where magnetisation transfer occurs faster. The pulse sequence is the same as PDSD but during t_{mix}, a weak field on the ^{1}H channel broadens the carbon lines and facilitates spin diffusion (Takegoshi et al., 2001) (see section 1.6.4 for a schematic illustration of the pulse sequence).

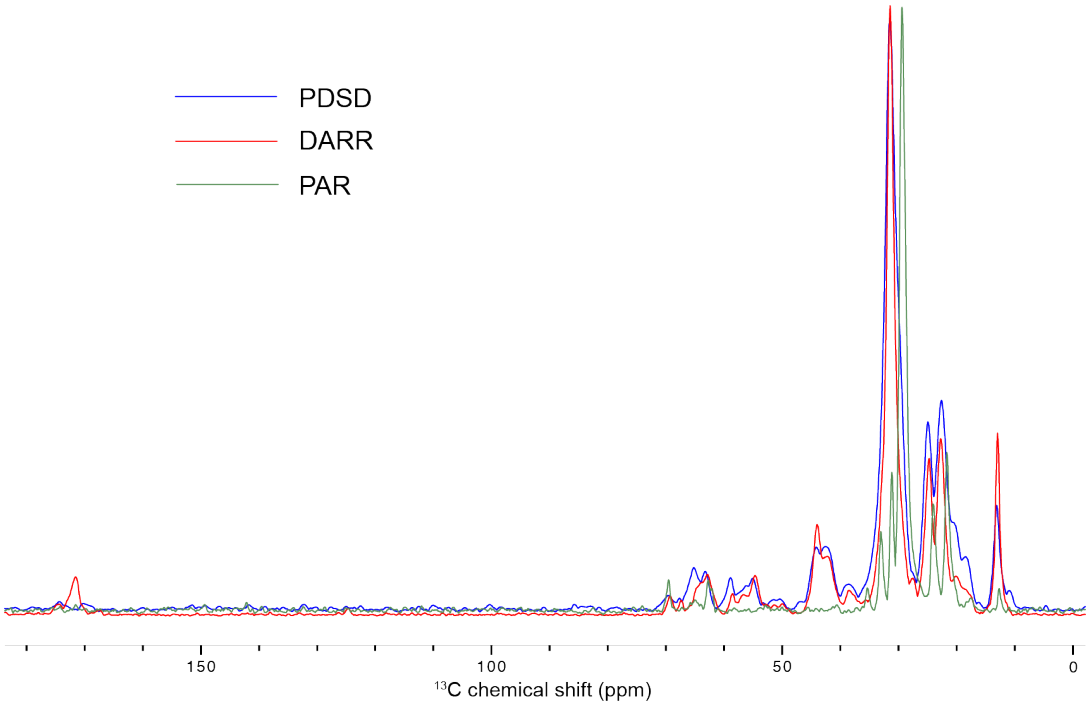


Figure 3.14: 1D ^{13}C-^{13}C spectra of Neu* [U–^{15}N, ^{13}C]-G_{665}/L_{668} labelled sample using various pulse sequences. Spectrum recorded at 850 MHz with 10 kHz MAS at -15 °C.

Proton-assisted recoupling (PAR)

PAR involves irradiation on both carbon and proton channels and avoids all rotary resonance and Hartmann-Hahn conditions (Lewandowski et al., 2009).

Figure 3.14 shows 1D projections of all of the above pulse sequences run on a U-^{13}C G_{665}/L_{668} Neu* sample. While peaks of significant resolution and intensity are seen

for both PDSD and DARR in the aliphatic region, the carbonyl region can only be observed in the DARR spectrum. This is in concordance with the information in scientific literature and that which has previously been acquired in our group. Consequently, all subsequent homonuclear 2D experiments were performed using the DARR experiment.

3.8.2 Resonance assignment

In order to assign resonances in the [U–^{15}N, ^{13}C]-G_{665}/L_{668} and [U–^{15}N, ^{13}C]-T_{662}/I_{659} samples of Neu*, the resonances observed in the spectrum were compared with published average chemical shift values for glycine, leucine, isoleucine and threonine found at the Biological Magnetic Resonance Databank (BMRB) (`http://www.bmrb.wisc.edu/`) (Ulrich et al., 2008). All peaks corresponding to side chain carbons of these ^{13}C-labelled amino acids in both samples of Neu* were successfully assigned (apart from L_{668} Cδ_1 & δ_2 and T_{662} Cα & Cβ) as they matched well with published values in BMRB. Even though the peaks were broad, the amino acids were selected to ensure minimal overlap and the ambiguities were observed only in the cases of T_{662} Cα & Cβ and L_{668} Cδ_1 & Cδ_2 which couldn't be resolved and were labelled collectively as C$\alpha\backslash\beta$ and C$\delta_1\backslash\delta_2$ respectively. Tables 3.3, 3.4, 3.5 and 3.6 list the ^{13}C Neu* chemical shifts and the BMRB values for all labelled amino acids used in this study. It is worth noting that the random-coil chemical shift values given in the BMRB database are referenced with respect to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and the Neu* experimental chemical shifts are referenced with respect to tetramethylsilane (TMS). Hence 2.0 ppm are subtracted from all entires in the BMRB before considering them in this study.

L_{668} chemical shifts (ppm)		
Nucleus	Neu* shift	BMRB
CO	176.0	175.3
Cα	56.13	53.95
Cβ	39.65	40.61
Cγ	24.30	24.78
Cδ_1	21.52	22.68
Cδ_2	21.52	22.07

Table 3.3: ^{13}C chemical shift data for L_{668}.

G_{665} chemical shifts (ppm)		
Nucleus	Neu* shift	BMRB
CO_1	172.7	171.9
CO_2	169.0	171.9
CO_3	171.3	171.9
Cα_1	45.76	43.36
Cα_2	43.23	43.36
Cα_3	50.79	43.36

Table 3.4: ^{13}C chemical shift data for G_{665}. Nuclei in a secondary environment are highlighted in purple and those in a tertiary environment are highlighted in red.

T_{662} chemical shifts (ppm)		
Nucleus	Neu* shift	BMRB
CO	174.1	172.6
Cα	65.60	60.27
Cβ	65.60	67.71
Cγ	19.7	19.56

Table 3.5: ^{13}C chemical shift data for T_{662}.

I_{659} chemical shifts (ppm)		
Nucleus	Neu* shift	BMRB
CO	175.3	173.93
Cα	64.21	59.68
Cβ	36.03	36.57
Cγ_1	27.48	25.74
Cγ_2	15.54	15.53
Cδ	12.35	11.4

Table 3.6: ^{13}C chemical shift data for I_{659}.

3.8.3 Secondary chemical shift analysis

The secondary chemical shift can be described as the deviation of the chemical shift of a nucleus from the random-coil shift. In particular, Cα, Cβ and CO nuclei form a useful probe into the secondary structure of proteins for carbon-detected ssNMR experiments (Spera and Bax, 1991; Wishart et al., 1991; Wishart and Sykes, 1994b,a). The secondary chemical shift, $\Delta\delta_S^i$, can be calculated using the following equation:

$$\Delta\delta_S^i = \Delta\delta_S^{obs} - \Delta\delta_S^{rc}$$

where,

$\Delta\delta_S^{obs}$ is the observed chemical shift and
$\Delta\delta_S^{rc}$ is the random-coil chemical shift.

In a study of 442 residues whose secondary structure is known with a high degree of precision, and Cα, Cβ and CO shifts are also known, it was found that for an α-helix the average Cα and Cβ secondary shifts are 3.09 ppm and -0.38 ppm respectively (Spera and Bax, 1991). In a similar study, the average secondary shift for CO was found to be 1.7 ppm downfield (Wishart and Sykes, 1994b). These values match well with the average values for the four labelled amino acids in our study: 3.62 ppm, -2.92 ppm and 1.09 ppm for Cα, Cβ and CO respectively. For each individual amino acid these values are illustrated in Figure 3.15 and the values are consistent with that of an α-helical structure and also in agreement with CD data.

As shall be explained in section 3.8.4, secondary and tertiary species were also observed for G_{665}. For both of these, the Cα secondary shifts are consistent with an α-helix but the CO shift values are negative and hence are not. The Cα shift values are a more reliable indicator of secondary structure than the CO shift values (Wang and Jardetzky, 2002). At the very least, however, it can be said that the results for these species are inconclusive in showing a conformational preference.

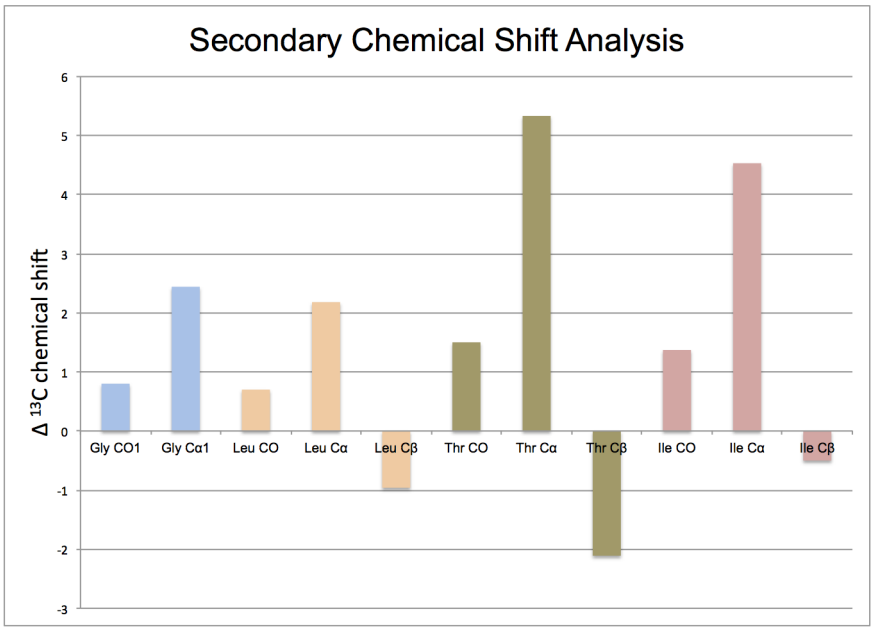


Figure 3.15: Secondary chemical shift analysis of all labelled amino acids in Neu* samples. Positive Cα and CO shifts and negative Cβ shifts are consistent with the α-helical structure of the peptide.

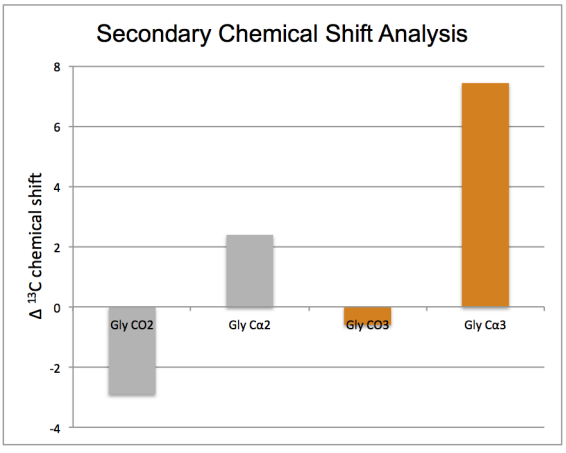


Figure 3.16: Secondary chemical shift analysis of secondary and tertiary species of G_{665}.

3.8.4 2D DARR spectra to investigate model A

After collecting an alanine CP spectrum for referencing/check and optimising CP on [U–^{15}N, ^{13}C]-G_{665}/L_{668} sample of Neu* in DMPC/cholesterol liposomes, 2D MAS ssNMR DARR experiments were performed. This would allow the observation of any inter- and intra-residue dipolar coupling for the Neu* dimer. As stated in section 3.3, should model A be representative of the dimeric interface of the Neu* dimer, it is expected that inter-helical polarisation transfer should be observed in the form of cross-peaks between [U–^{15}N, ^{13}C]-G_{665} and [U–^{15}N, ^{13}C]-L_{668} in the DARR spectrum.

Long (400 ms) mixing time

A 2D ^{13}C–^{13}C DARR spectrum obtained using a 400 ms mixing time is shown in Figure 3.17. Several cross peaks on either side of the diagonal were observed. The strong signals on the diagonal were due to either the labelled amino acids in Neu* or from natural abundance ^{13}C present in DMPC lipids. Since only one amino acid was labelled per peptide, and because the labelling of amino acids was done such that the average chemical shift of each resonance was as different as possible, peaks were largely well resolved in both the carbonyl and aliphatic regions. The only exceptions to this were Cδ_1 and Cδ_2 of [U–^{15}N, ^{13}C]-L_{668} which were unresolvable due their chemical shifts being very similar and the broadness of the peaks.

At a mixing time of 400 ms, several intra-residue connectivities were observed. For [U–^{15}N, ^{13}C]-L_{668}, intra-residue cross peaks were observed between Cα–Cβ, Cα–C$\delta_{1/2}$, Cα–Cγ, Cα–CO, Cβ–C$\delta_{1/2}$, Cβ–Cγ, Cβ–CO, Cγ–CO and C$\delta_{1/2}$–CO. For G_{665}, the only intra-residue Cα–CO cross peak was observed on both sides of the diagonal. However, it is observed in 3 different environments: one major species Cα_1–CO_1 and two minor species Cα_2–CO_2 &C α_3–CO_3.

Several peaks which can be attributed to inter-helical magnetisation transfer are also observed:

$$L_{668}C\beta\text{–}G_{665}CO$$

$$L_{668}C\gamma\text{–}G_{665}C\alpha_2$$

$$L_{668}C\gamma\text{–}G_{665}C\alpha_3$$

$$L_{668}C\gamma\text{–}G_{665}CO.$$

These are highlighted in green in Figure 3.17. In the aliphatic region, both of the minor environments of glycine form $L_{668}C\gamma$–$G_{665}C\alpha$ inter-helical cross peaks. Interestingly, cross peaks between the minor environments are also observed in the spectrum which may indicate a slow conformational exchange (relative to NMR timescales) yielding two separate peaks in the NMR spectrum. These are shown in purple in Figure 3.17.

The major species is responsible for the inter-helical $L_{668}C\beta$–$G_{665}CO$ cross peak on one side of the diagonal and $L_{668}C\gamma$–$G_{665}CO$ cross peak on the other. Such asymmetry can be ascribed to non-uniform CP signal enhancement as well as some inherent features of DARR recoupling (Ohashi and Takegoshi, 2006).

The presence of these inter-helical cross peaks is of great significance. Firstly, it shows that the sample preparation method has yielded a sample in which Neu* is not only reconstituted with an α-helical conformation in DMPC/cholesterol liposomes but also, crucially, that it dimerises in the hydrated lipid bilayer environment. Furthermore, the results support model A for the dimerisation of Neu* as they indicate that L_{668} and G_{665} are close in space. This in turn supports the hypothesis that the $A_{661}XXXG_{665}$ motif stabilses the dimer.

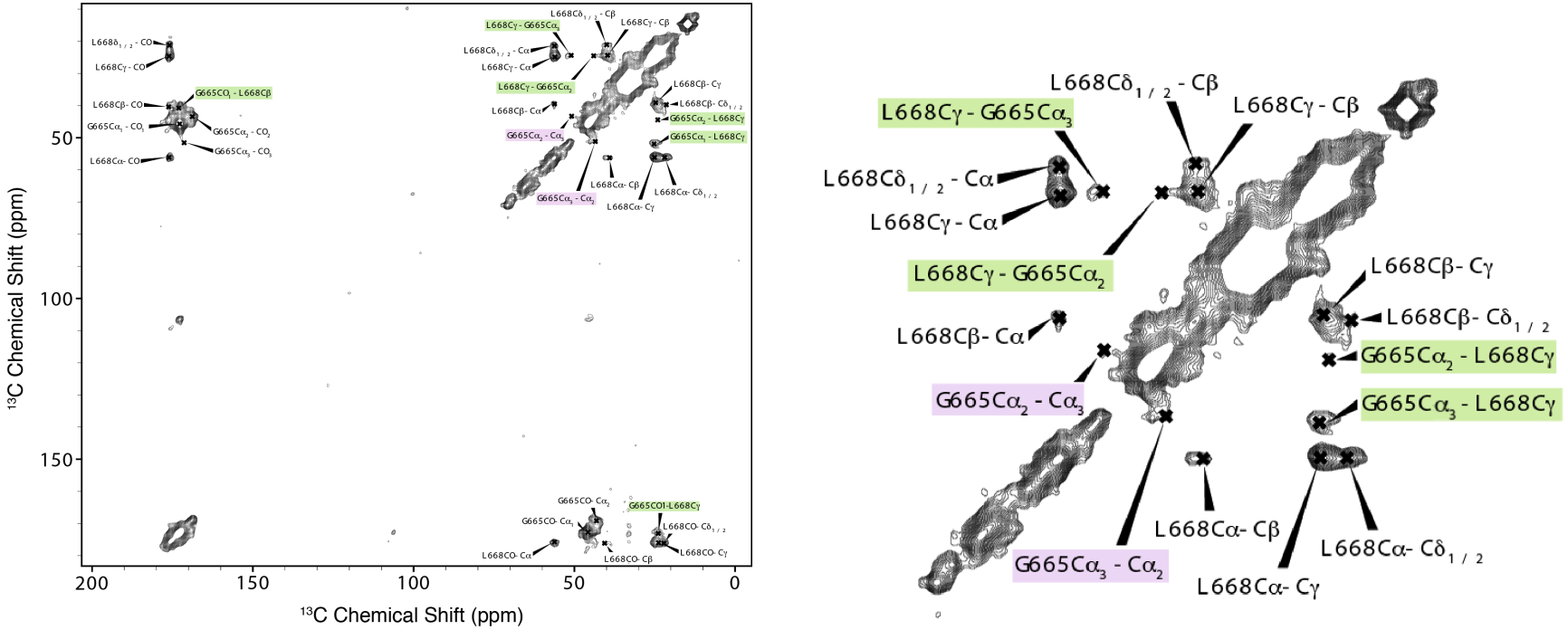


Figure 3.17: 2D spinning frequency ssNMR ^{13}C–^{13}C DARR correlation spectrum of [U–^{15}N, ^{13}C]-G_{665}/L_{668} labelled Neu* in DMPC/cholesterol liposomes at a mixing time of 400 ms. Spectrum was acquired at 600 MHz with 10 kHz MAS at -15 °C. The complete spectrum is shown on the left whereas the aliphatic region is expanded on the right. All inter-helical cross peaks are highlighted in green whereas cross peaks which are thought to arise from putative conformational exchange are labelled in purple.

Short (30 ms) mixing time

A short mixing time (30 ms) DARR experiment was also performed under the same conditions as above. The assignment of intra-residue cross peaks obtained in the long mixing time experiment were confirmed. Some of the cross peaks disappeared because, at short mixing times, long range through-space couplings cannot be observed. But more surprisingly, at least three inter-helical cross peaks are still clearly observed (highlighted in green in Figure 3.18), though just on one side of the diagonal. This could indicate very short intermolecular distances as a result of strong dimerisation. Furthermore, the cross peaks between the two minor environments of glycine are again observed (highlighted in purple in Figure 3.18).

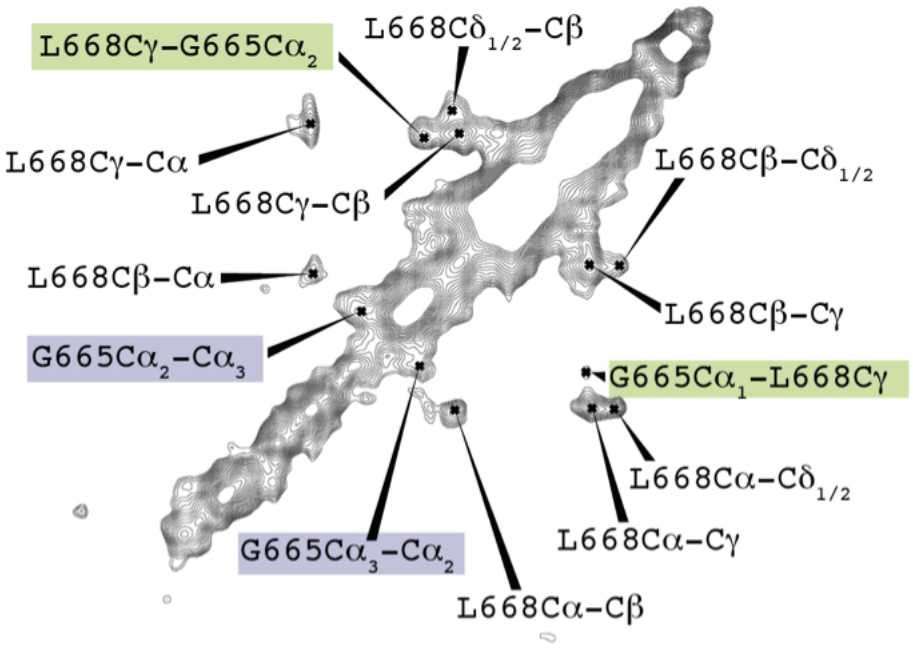


Figure 3.18: 2D MAS ssNMR ^{13}C–^{13}C DARR correlation spectrum of the aliphatic region of [U–^{15}N, ^{13}C]-G_{665}/L_{668} labelled Neu* in DMPC/cholesterol liposomes at a mixing time of 30 ms. Spectrum was acquired at 600 MHz with 10 kHz spinning frequency at -15 °C.

3.8.5 2D DARR spectra to investigate model B

As above, after collecting an alanine CP spectrum for referencing/check and optimising CP on a [U–^{15}N, ^{13}C]-T_{662}/I_{659} sample of Neu* in DMPC/cholesterol liposomes,

2D MAS ssNMR DARR experiments were performed. As stated in section 3.3, should model B be representative of the dimeric interface of the Neu* dimer, it is expected that inter-helical polarisation transfer should be observed in the form of cross-peaks in the DARR spectrum.

Long (400 ms) mixing time

A 2D ^{13}C–^{13}C DARR spectrum obtained using a 400 ms mixing time is shown in Figure 3.19. Several cross peaks on either side of the diagonal are observed. As above, the strong signals on the diagonal are due to either the labelled amino acids in Neu* or from natural abundance ^{13}C present in DMPC lipids. T_{662} Cα & Cβ were unresolvable due their chemical shifts being very similar and the broadness of the peaks.

At a mixing time of 400 ms, several intra-residue connectivities are observed. For I_{659}, intra-residue cross peaks are observed between Cα–Cβ, Cα–Cγ_1, Cα–Cγ_2, Cα–Cδ, Cα–CO, Cβ–Cγ_1, Cβ–Cγ_2, Cβ–Cδ, Cβ–CO, Cγ_1–Cγ_2, Cγ_1–Cδ, Cγ_2–CO and Cδ–CO. For T_{662}, intra-residue cross peaks are observed between α & Cβ—Cγ, Cα & Cβ–CO and Cγ–CO.

No evidence of any inter-helical magnetisation transfer was observed and as such the results here do not support model B which states that the $I_{659}XXXV_{663}$ motif in Neu* stabilises the dimer. However, 'splitting of the peaks' is observed for a number of peaks in I_{659} which may suggests a high degree of mobility of the peptides and possibly conformational exchange between multiple states.

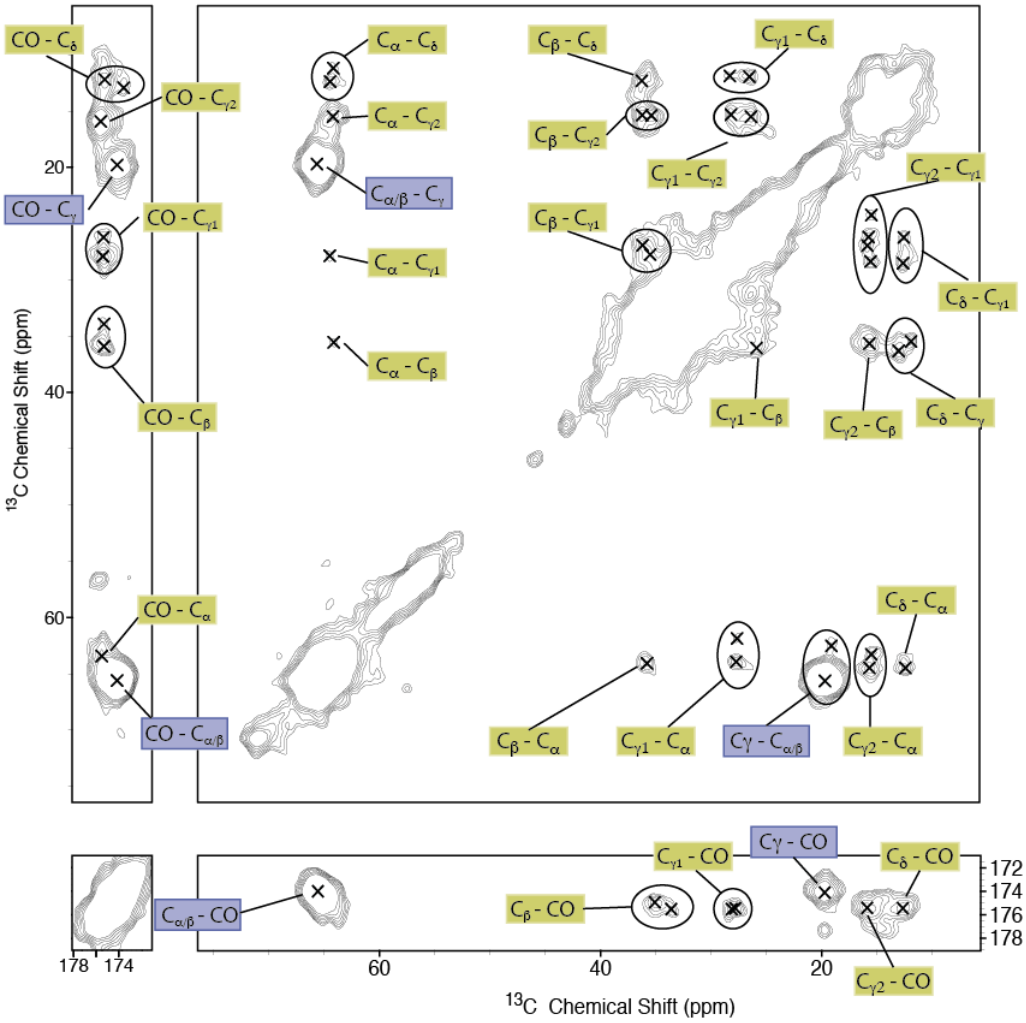


Figure 3.19: 2D MAS ssNMR ^{13}C–^{13}C DARR correlation spectrum of the aliphatic region of [U–^{15}N, ^{13}C]-T_{662}/I_{659} labelled Neu* in DMPC/cholesterol liposomes at a mixing time of 400 ms. All intra-residue T_{662} cross peaks are highlighted in purple and intra-residue I_{659} cross peaks are highlighted in mustard. Spectrum was acquired at 600 MHz with 10 kHz MAS at -15 °C.

3.9 Summary and Conclusion

Work presented in this chapter sought to investigate the structure and stabilising interactions of the Neu* dimer. The ssNMR data presented was obtained from the Neu* protein TM dimer in a hydrated lipid bilayer containing 10% cholesterol. The rigidity of the membrane can be controlled through the addition of cholesterol, also providing it with a 'native-like' heterogeneity. A pair of peptides were designed to investigate which of the models for the Neu* TM can be corroborated, one for model

A and one for model B. Upon receipt from the manufacturer, these were purified and reconstituted into DMPC liposomes.

Circular dichroism indicated that these adopt an α-helical secondary structure within the membrane and this was confirmed through secondary chemical shift analysis of the labelled nuclei in the peptides.

For each sample, conditions such as CP contact time and various experimental parameters were optimised. Following this, 2D MAS ssNMR ^{13}C–^{13}C DARR correlation spectra were acquired to check for any inter-residue correlations which may strengthen the hypothesis behind either model A or model B. For model A, inter-helical correlations were observed at both long and short mixing times indicating that the residues were extremely close in space to the hypothesised dimeric interface. For model B, no such correlations were observed an no conclusions regarding spatial proximity of the labelled residues could be drawn.

However, both models showed evidence of existence of multiple conformational states and/or exchange between them. For model A, G_{665} was observed in 3 different environments and 2 of those had cross-peaks between the nuclei in different states which could be taken as evidence of exchange between them. For model B, various nuclei in I_{659} exhibited 'splitting of peaks' which could once again be taken as evidence of a high degree of mobility. Evidence of more than one stable conformation supports the theory that TM dimers can exist in 'active' and 'inactive' states and switch between the conformations (introduced in section 1.2).

Previous MAS ssNMR analyses performed on this set of peptides were performed on a sample prepared using a different protocol involving freeze-drying the sample. Also, the protein : cholesterol : DMPC ratio was 1 : 5 : 15 compared to 1 : 2 : 20 here (Ellis et al., unpublished data). Even though spectra with good signal to noise were obtained in the previous study, no cross-helical correlations were observed. The existence of those peaks in this study highlights the importance of cholesterol on controlling the rigidity of the membrane and also reaffirms the importance of avoiding placing the peptides in conditions which may compromise their structure.

Chapter 4

Cloning and Expression of PDGFβR transmembrane domain

4.1 Introduction

Platelet-derived growth factor-β receptor (PDGFβR) is a type III cell-surface receptor tyrosine kinase (RTK) that plays an essential role in cellular processes such as mitogenesis, chemotaxis, angiogenesis and proliferation (Heldin and Westermark, 1999). PDGFβR activation is induced when a ligand binds to the ectodomains of the PDGFβR dimer and subsequently results in the autophosphorylation of the intracellular kinase domains (see section 1.2) for details).

Much of the work towards the high-resolution structural characterisation of PDGFβR has centred around the soluble domains which are amenable to crystallographic studies (Shim et al., 2010; Yang et al., 2008). More recent *in vivo* and *in vitro* biophysical studies have demonstrated the ability of the transmembrane (TM) domain to self-associate with the juxtamembrane (JM) regions playing an important role in oligomerisation (Muhle-Goll et al., 2012; Oates et al., 2010).

The role of the TM domain in mediating the dimerisation and activation of the receptor is of significant interest also because the smallest known oncoprotein, E5 from bovine papillomavirus type-1, is known to induce ligand-independent activa-

tion of the receptor through a mechanism which is not yet properly understood. It is thought that the self-assembly of E5 into homodimers and its subsequent interaction with the TM domain of PDGFβR triggers constitutive activation of the receptor.

It is known that E5 must exist as a dimer in order to bind and activate PDGFβR (Horwitz et al., 1988; Nilson et al., 1995). Mutagenesis and computational data suggest that the dimerisation of E5 generates two identical faces, each of which is able to bind a monomer of PDGFβR (Petti et al., 1997; Surti et al., 1998). The specific interactions between E5 and PDGFβR which stabilise the complex involve hydrogen-bonding between Q_{17} of E5 and T_{545} of PDGFβR, and electrostatic interactions between D_{33} of E5 and K_{531} of PDGFβR (Talbert-Slagle and DiMaio, 2009). This is illustrated in Figure 4.1

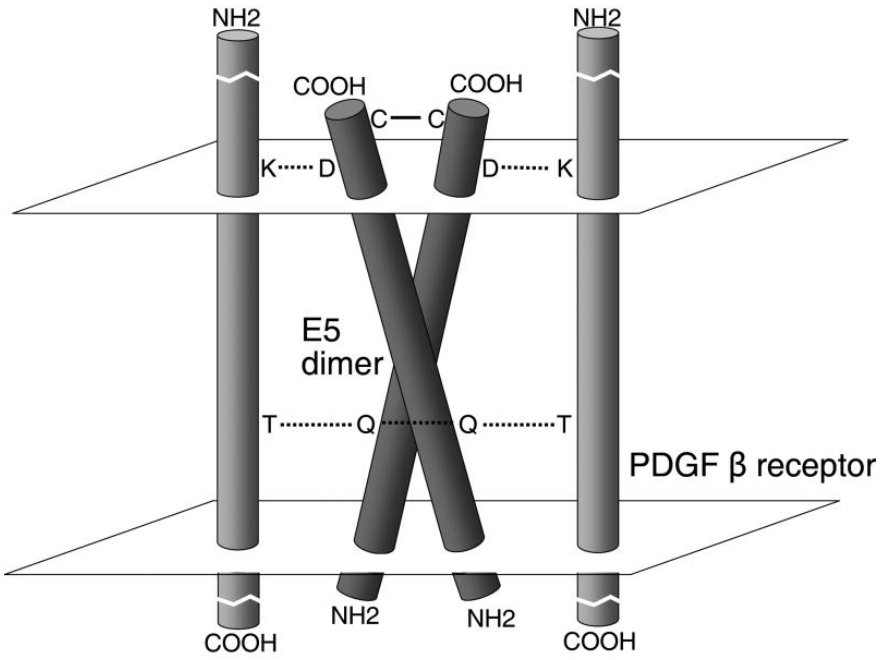


Figure 4.1: Schematic of the E5/PDGFβR complex. Adapted from (Talbert-Slagle and DiMaio, 2009).

The aim in this study was to develop a protocol for the heterologous expression of the transmembrane domain of PDGFβR (PDGFβR-TM) for structural analyses using solid- and solution-state NMR. High yield production of the peptide would allow us to obtain currently missing direct evidence of E5/PDGFβR binding in membrane

mimetics through chemical shift perturbations in NMR spectra. Furthermore, the effect of E5 binding on juxtamembrane regions was also be investigated by inclusion of those regions in the protein construct.

4.2 Initial Expression Trials

4.2.1 Choice of plasmid and cell lines

The triple resonance experiments conducted in this study required that an efficient protocol be devised for expressing high yields of PDGFβR-TM. This is primarily because high costs of isotopic labels can render the study infeasible if good yields are not obtained. A single NMR sample of 180 μL requires a protein concentration of at least 0.3 mM for 3D analyses and it is of utmost importance to maximise yields to make it economically feasible.

Previous work by Dr Joanne Oates to characterise the role of TM and JM domains in oligomerisation of PDGFβR involved the construction of plasmids where PDGFβR-TM with either 20, 30 or 40 residues on either side were cloned into a pET30a vector (Novagen) for expression in *E. coli* (Oates et al., 2010). All of these vectors contained a His-tag which allowed for purification using IMAC. These constructs will henceforth be referred to as pET30a-PDGFβR-TM-20, pET30a-PDGFβR-TM-30 and pET30a-PDGFβR-TM-40 respectively. It was thought that these constructs would form an ideal starting point for the expression of PDGFβR-TM for NMR analyses.

Cells were grown directly from glycerol stocks which had the constructs of interest in 3 different types of *E. coli* cell lines: BL21(DE3)pLysS; C41; and C43. BL21(DE3)pLysS cells allow protein expression of any gene that is under the control of a T7 promoter and has a ribosome binding site. These cells also contain a pLysS plasmid which carries the gene encoding T7 lysozyme which lowers 'leaky' expression but does not interfere with the level of expression achieved following induction by isopropyl β-D-1-thiogalactopyranoside (IPTG). C41 cells are derived from BL21(DE3),

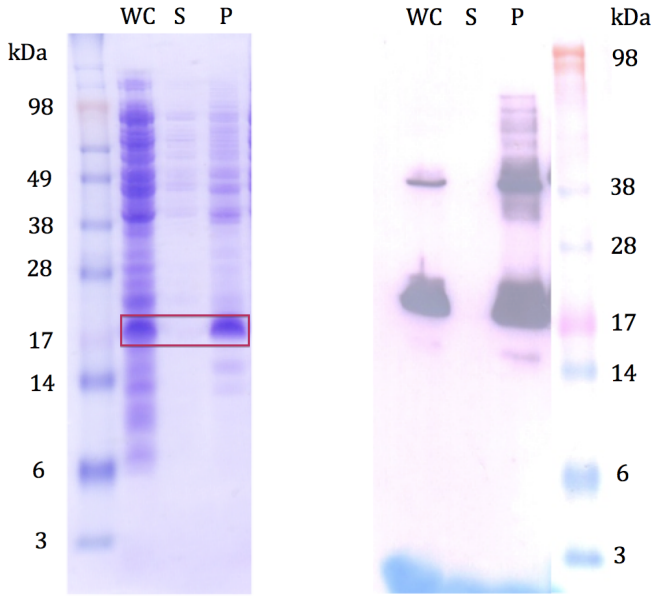
and contain a mutation which prevents cell death associated with recombinant expression of some proteins which may be toxic to the cells. C43 cells are a derivative of C41 which can express a different set of toxic proteins than C41.

The C41 and C43 cells did not grow in LB medium in the presence of antibiotics even after several attempts. The transformed BL21(DE3)pLysS cells grew in the presence of antibiotics. The plasmid was extracted and sent for sequencing to verify that no mutations occurred in the sequence.

4.2.2 Expression of pET30a-PDGFβR-TM-40

Prior to optimising conditions for expression in minimal media, the pET30a-PDGFβR-TM-40 plasmid was transformed into BL21(DE3)pLysS cells (see section 2.5.3) and expressed in LB (see section 2.6.1) as per existing protocols (Oates et al., 2010) to check for expression levels. Once the cells were harvested, a portion of the whole cells (WC) were lysed and centrifuged. The resulting supernatant (S) and pellet (P), along with the whole cells were analysed on an SDS-PAGE gel as well as a western blot as detailed in sections 2.8.1 and 2.8.4. The expected molecular mass of the expressed protein is 18.1 KDa and the gel (Figure 4.2a) revealed a thick band at the position of the expected monomer in the whole cell as well as pellet fractions (highlighted in red box in panel (a)). This suggested that the protein was being expressed in inclusion bodies. Inclusion bodies are formed for 70% of recombinant proteins over-expressed in *E. coli* (Yang et al., 2011). This is usually due to the lack of eukaryotic chaperones and the saturation of *E. coli* post-translational machinery (Singh et al., 2015).

The western blot (Figure 4.2b) corroborates that the bands observed are of PDGFβR-TM-40 as only proteins with a His-tag will show up on the blot. A dimer is also observed at about 40 kDa. This is slightly more than the expected 36.2 kDa but such anomalous migration has frequently been reported in literature for membrane proteins, especially for higher order oligomers (Rath et al., 2009). In particular, slow migration of higher order oligomers of PDGFβR-TM has been observed

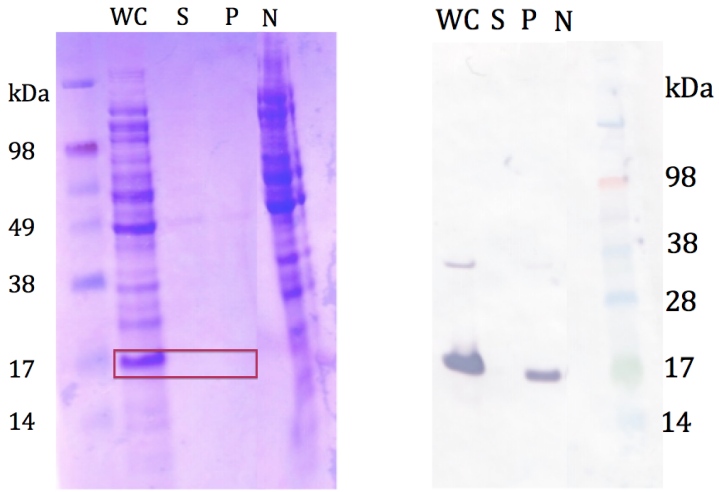


(a) SDS-PAGE **(b)** Western blot

Figure 4.2: Expression pET30a-PDGFβR-TM-40 in LB. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. Red box highlights the position of the expected monomer PDGFβR-TM-40 band. WC: whole cells, S: supernatant after lysis and centrifugation of whole cells, P: pellet after lysis and centrifugation of whole cells.

previously (Oates et al., 2010).

Once good expression levels were observed in rich medium, the PDGFβR-TM-40 protein was expressed in M9 minimal media. The SDS-PAGE gel shown in Figure 4.3a reveals that even though the expression levels drop considerably in minimal medium, clear bands corresponding to the protein were seen (highlighted in red box in panel (a)). A anti-His western blot confirmed this and a dimer of the protein was once again observed (Figure 4.3b). A lane of non-induced whole cells (N) was run on both SDS-PAGE and western blots to ensure there was no 'leaky' expression and that the cells weren't natively producing a protein which was being bound by anti-His antibody during western blotting.



(a) SDS-PAGE **(b)** Western blot

Figure 4.3: Expression pET30a-PDGFβR-TM-40 in M9 minimal media. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. Red box highlights the position of the expected monomer PDGFβR-TM-40 band. WC: whole cells, S: supernatant after lysis and centrifugation of whole cells, P: pellet after lysis and centrifugation of whole cells, N: non-induced whole cells.

4.2.3 Comparison of expressions with varying JM lengths

As mentioned previously, the juxtamembrane regions are known to play an important role in the activation and oligomerisation of the PDGFβR. It was observed by Oates and coworkers (Oates et al., 2010) that the PDGFβR-TM on its own forms higher order oligomers *in vitro* which are not biologically relevant. It was further observed that upon the addition of 40 juxtamembrane residues on either side, the peptide existed exclusively as monomers and dimers.

Since the goal for this study was to prepare samples for NMR analyses, it was preferred that the number of JM amino acids to be included in the sample be minimised so as to avoid spectral crowding. This is especially important in the case of a membrane protein embedded in detergent micelles which can lead to broad peaks in an NMR spectrum. Figure 4.4 shows a western blot of the expression levels of pET30a-PDGFβR-TM-40 (JM40), pET30a-PDGFβR-TM-30 (JM30) and pET30a-PDGFβR-TM-20 (JM20) in M9 minimal medium. Whole cells (WC) as well as pellets (P) after lysis and centrifugation of whole cells were analysed. Good expres-

sion levels were observed for the smallest protein, PDGFβR-TM with 20 resides on either side of the TM. For these reasons, all further work was carried out using this construct.

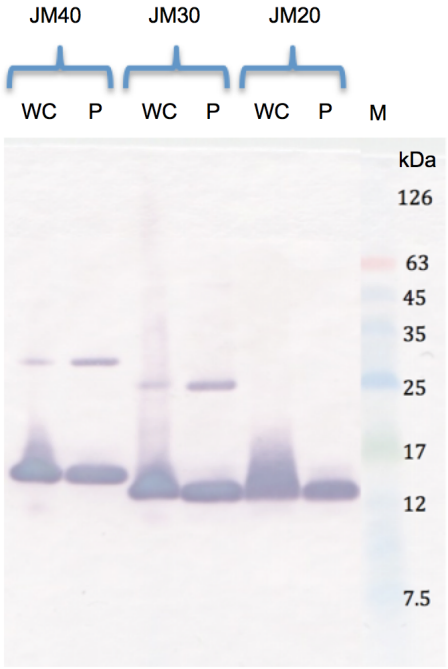


Figure 4.4: Expression levels of pET30a-PDGFβR-TM with varying JM lengths observed using an anti-His western blot. Cells were grown at 37 ° in M9 medium for 4 hours after induction with 1 mM IPTG. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, M: protein ladder, JMX: PDGFβR-TM with X residues on either side.

4.3 Optimisation of Protein Yield

4.3.1 Glucose concentration

When preparing isotopically labelled protein for NMR analyses, it is of critical importance that expensive labels be used most efficiently. Studies have shown that protein yields can increase several fold by increasing the concentration of glucose in M9 medium to up to 4 g/L. However, beyond that point, even though there is some increase in protein yield, it is not economically feasible (Berthold et al., 2011; Marley et al., 2001). Figure 4.5 shows that a significant improvement in expression levels was

obtained after doubling the glucose concentration from 2 g/L to 4 g/L (highlighted in red box). Hence 4 g/L of glucose was used in subsequent expressions.

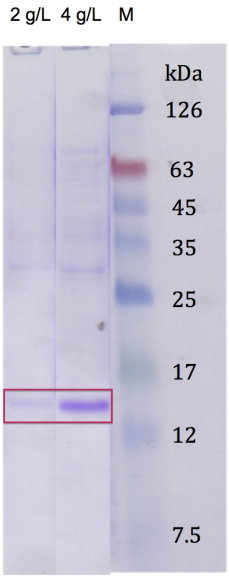


Figure 4.5: Expression levels of pET30a-PDGFβR-TM-20 with varying glucose concentrations. Cells were grown at 37 °C for 4 hours after induction with 1 mM IPTG. Red box highlights the position of the expected monomer PDGFβR-TM-20 band. Lanes labelled with the concentration of glucose used in expressing the protein, M: protein ladder

4.3.2 Addition of Basal Medium Eagle (BME) vitamins

Supplements to the M9 minimal medium such as vitamins have also been known to enhance the growth and expression of *E. coli* cells (Jansson et al., 1996; Kainosho, 1997; Cai et al., 1998). For PDGFβR-TM-20 expression, Basal Medium Eagle (BME) vitamins (Sigma) were chosen and consist of vitamins such as thiamine, riboflavin and folic acid to stimulate cell growth. Figure 4.6 shows an SDS-PAGE gel which reveals that the addition of vitamins did not lead to any improvement in expression levels.

4.3.3 Addition of iron (III) chloride

It has been reported by Berthold and coworkers (Berthold et al., 2011) that addition of a trace metal element solution to media improves the yield of the expressed

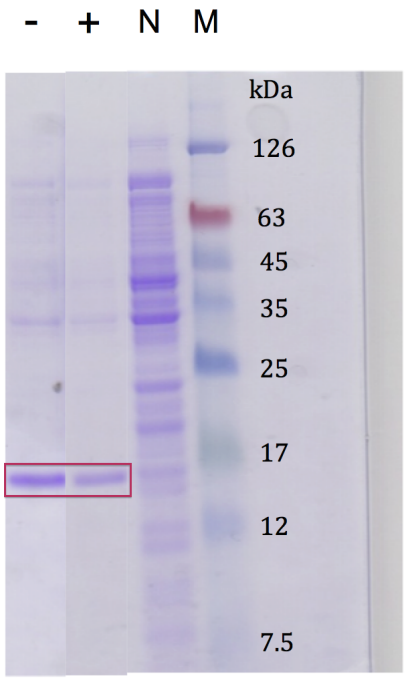


Figure 4.6: Expression levels of pET30a-PDGFβR-TM-20 with and without the addition of vitamins to the M9 minimal medium observed on an SDS-PAGE gel. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. Red box highlights the position of the expected monomer PDGFβR-TM-20 band. Lane marked (-) represents the expression without vitamins and (+) represents the expression with BME vitamins, M: protein ladder.

protein. The effect of trace metal element solution has also been studied by Studier (Studier, 2005), where they showed that the addition of 100 μL of 100 mM $FeCl_3$ is enough to get very close to the maximal yield that is obtained by adding a complete trace metal element solution. However, as can be seen in Figure 4.7, no improvement in expression level was observed upon the addition of iron (III) ions.

4.3.4 Effect of temperature variation

It has been reported in scientific literature that if the rate of transcription is coupled with that of translation and that of post-translational events then that can significantly increase protein yield (Berthold et al., 2011). In order to enable this to happen, the rates of cellular metabolism can be affected by switching to a weaker promoter (such as T5 or arabinose) or by using *E. coli* strains which overcome the toxicity associated with the overexpression of the T7 promoter. However, the sim-

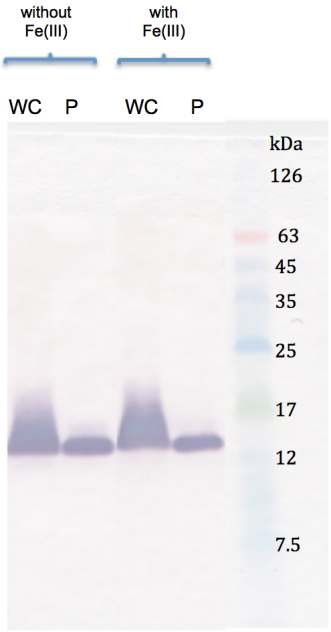


Figure 4.7: Expression levels of pET30a-PDGFβR-TM-20 with and without the addition of iron (III) to M9 minimal medium observed on an anti-His western blot. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, M: protein ladder.

plest way of achieving optimal rates of cellular metabolism is to scan a range of temperatures. Figure 4.8 shows an SDS-PAGE gel where pET30a-PDGFβR-TM-20 was expressed at 37 °C, 30 °C, 22 °C and 15 °C for 4 hours, 6 hours, 12 hours and 18 hours respectively. After harvesting cells at each temperature, some of the whole cells (WC) were lysed and centrifuged. The resulting supernatant (S) and pellet (P) were also run on the gel. The pellet contains the inclusion bodies. It can be observed by looking at the pellet lanes at each temperature that the highest expression level (highlighted in a red box) is obtained at 37 °C, and it decreases as the temperature is reduced (very little protein being expressed at 15 °C). It was decided to perform all future expression of this protein at 37 °C.

4.3.5 IPTG concentration

Upon addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a bacterial culture, it initiates the transcription of the *lac* operon and hence induces recombinant protein expression when the gene is under the control of the lac operator. The final

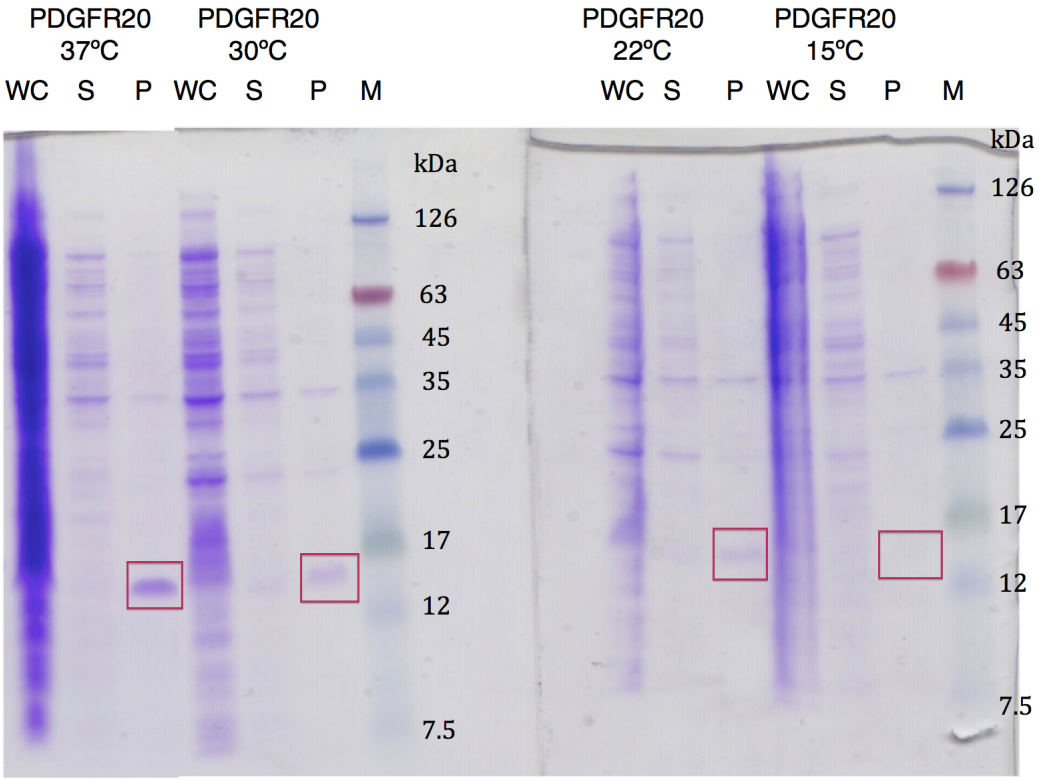


Figure 4.8: Expression levels of pET30a-PDGFβR-TM-20 at varying temperatures as observed on an SDS-PAGE gel. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, S: supernatant after lysis and centrifugation of whole cells, M: protein ladder.

concentration of IPTG used to induce *lac* repressor-regulated promoters can have a significant effect on the protein yield. With certain genes, slowly inducing transcription (with low IPTG concentrations) may be required to reduce the effects of certain toxic genes (Green and Sambrook, 2012). Figure 4.9 shows an SDS-PAGE gel where pET30a-PDGFβR-TM-20 was expressed with IPTG concentrations of 0.1 mM, 0.5 mM, 1 mM and 1.5 mM. The bands corresponding to the molecular weight of the protein have been highlighted in a red box. It was found that 1 mM yielded the strongest protein band in the whole cell and pellet fractions and all future expressions were performed with this IPTG concentration.

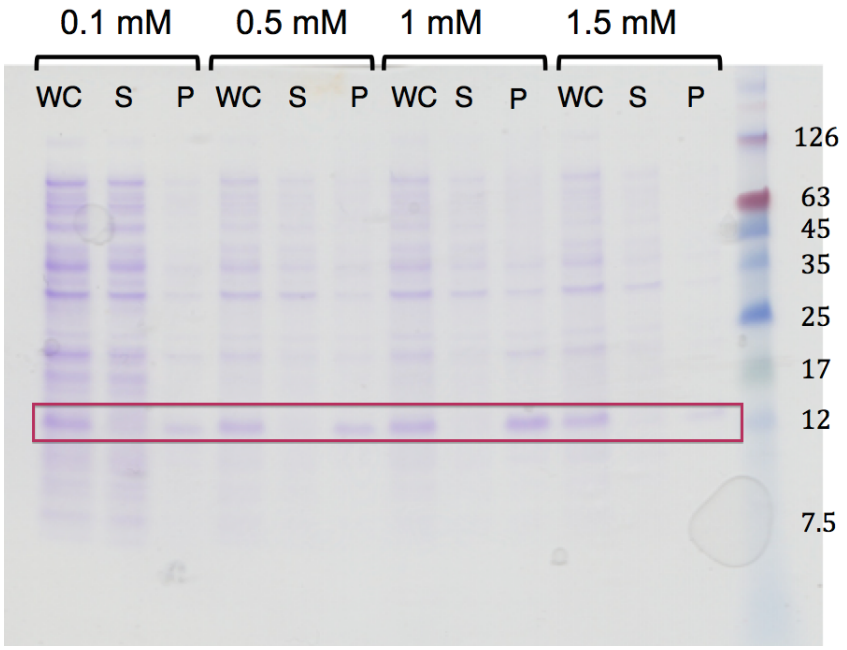
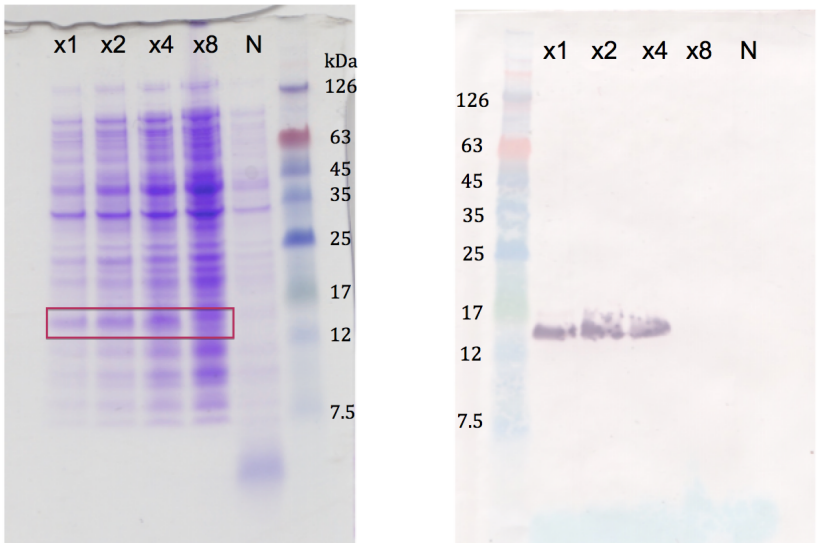


Figure 4.9: Expression levels of pET30a-PDGFβR-TM-20 at varying IPTG concentrations as observed on an SDS-PAGE gel. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, S: supernatant after lysis and centrifugation of whole cells

4.3.6 Media switching

Slower cell growth rates are observed when expressing recombinant proteins in M9 minimal medium as compared to LB, because the former contains a minimal number nutrients required for cells to grow. This results in reduced expression levels due to the cytotoxic effects associated with the plasmid exerting selective pressure (Baneyx, 1999). To minimise these effects, many studies have adopted the approach of transforming, inoculating and growing cells in rich medium and then subsequently switching to M9 minimal medium prior to inducing expression (Marley et al., 2001; Cai et al., 1998; Sivashanmugam et al., 2009). This method capitalises on the fast doubling rates of *E. coli* cells in rich media and hence greatly improves the efficiency with which high levels of recombinant protein can be obtained. This also removes unlabelled metabolites that can hinder growth and expression. Bacterial culture is gently pelleted once it reaches the mid-log phase, and then washed with minimal media without isotopic labels. It is then allowed to equilibrate in the new media (usually for one hour) before induction.



(a) SDS-PAGE **(b)** Western blot

Figure 4.10: Expression levels of pET30a-PDGFβR-TM-20 after concentrating the bacterial suspension to varying degrees. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG following concentration. ×1, ×2, ×4 and ×8 represent lanes where the the culture has been concentrated to 1, 2, 4 and 8 times their original concentration, respectively. N: non-induced whole cell.

4.3.7 High-cell-density expression

Further to switching the media prior to induction, it has also been observed that using minimal media with labelled isotopes to resuspend the cell pellet in a fraction of the original volume of rich media can lead to significant improvement in protein yield (Sivashanmugam et al., 2009). Figure 4.10a shows an SDS-PAGE gel of pET30a-PDGFβR-TM-20 expression after concentrating the bacterial suspension to varying degrees prior to induction: ×1, ×2, ×4 and ×8. Improvements in expression level were seen after doubling and quadrupling the concentration factor, but once it was raised to ×8 the protein did not express as can be seen in the blot in Figure 4.10b.

4.4 TOPO cloning of pET151-PDGFβR-TM-20

As is explained in detail in section 5.2.4, the pET30a-PDGFβR-TM-20 construct proved to be infeasible for the production of high yields of PDGFβR-TM-20. This was primarily due to two issues. Firstly, pET30a has an enterokinase cleavage site to remove an N-terminal thrombin site, S-tag and His-tag. Low efficiency cleavage meant cleaved protein could not be produced in a cost-effective manner for NMR structural studies. Secondly, the construct also encoded for a C-terminal His-tag which could not be cleaved; this can be problematic as the His-tag is can potentially lead to non-preferential interactions and promote aggregation (Hammarström et al., 2002). In light of this, it was decided that a new construct be made in the lab with a tobacco etch virus (TEV) protease cleavage site.

4.4.1 TOPO cloning

TOPO cloning is a molecular biology technique which relies on the inherent biological activity of DNA topoisomerase I. PCR products are cloned such that they have CACC added at the 5' end. As illustrated in Figure 4.11, the overhang in the cloning vector (GTGG) covalently reacts with the 5' end of the PCR product by annealing to the added bases and hence inserting the PCR product in the correct orientation.

PCR products were generated for the PDGFβR transmembrane domain and 20 juxtamembrane amino acids on either side (PDGFβR-TM-20) as detailed in section 2.3.3. These were analysed using agarose gel electrophoresis (secion 2.3.4) as shown in Figure 4.12. The resulting PCR product was observed as a band at $\sim$300 bp. The gene was then purified using a gel extraction kit (section 2.3.5). The TOPO reaction was subsequently performed by mixing the pET151 vector and the fresh PCR product as per manufacturer's instructions (Life Technologies, UK).

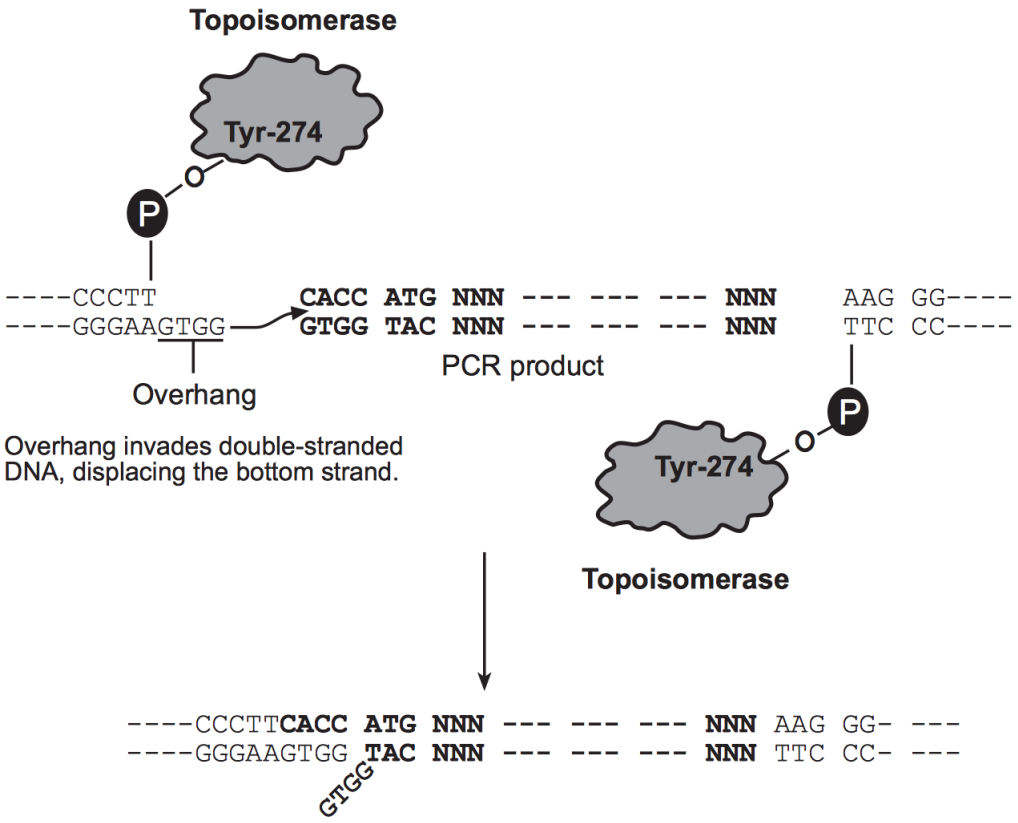


Figure 4.11: TOPO cloning schematic. Adapted from (Life Technologies pET Directional TOPO Manual, 2010).

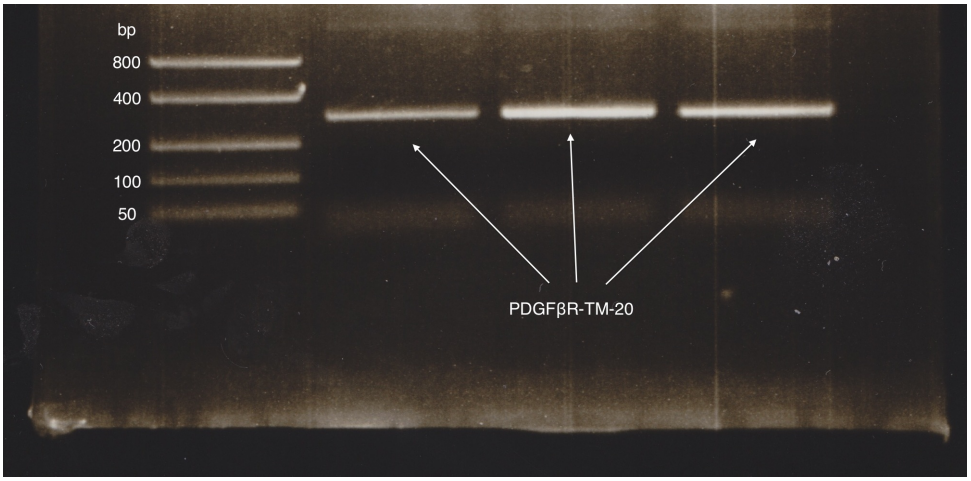


Figure 4.12: PCR products from the amplification of PDGFβR-TM-20 encoding with CACC added at the 5' for TOPO cloning.

4.4.2 Colony PCR

Colony PCR is a useful and convenient high-throughput method for establishing the presence or absence of insert DNA in plasmid constructs. Chemically competent cells are transformed using the TOPO cloning reaction directly as described in section 2.5.3. Several colonies were subsequently observed on the plate and Colony PCR (section 2.3.8) was performed on 28 of them. Figure 4.13 shows the products of the Colony PCR after being run on an agarose gel. Six colonies where there is clear indication of insert DNA being amplified (marked with red arrows in Figure 4.13) are grown in LB, their plasmids extracted and sent for sequencing. It was found that out of the six constructs sent, only one of them had the insert in the right way round. The DNA from the correct construct was saved for all future transformations.

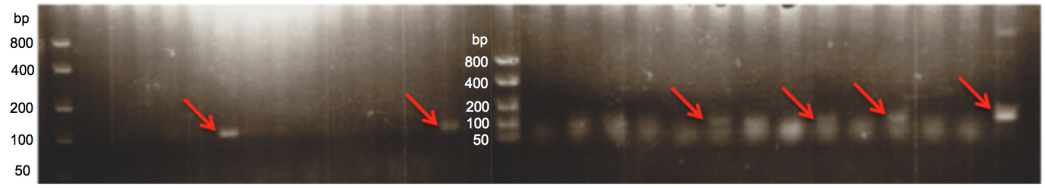
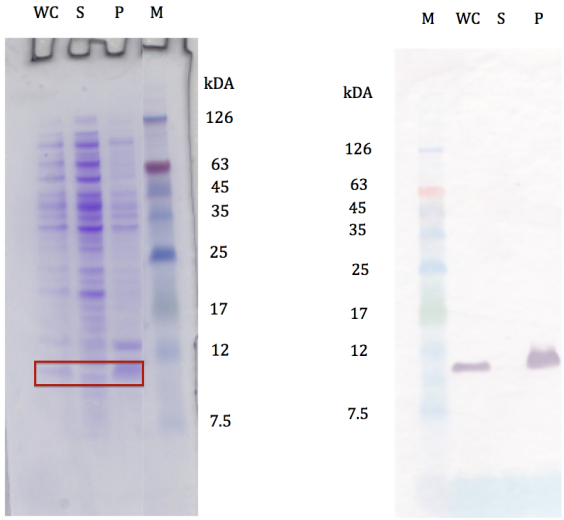


Figure 4.13: Colony PCR of the transformants after TOPO cloning. Colonies where evidence if DNA amplification was seen are marked with red arrows

4.4.3 Expression checks for PDGFβR-TM-20 and TEV protease

PDGFβR-TM-20

After successfully cloning the pET151-PDGFβR-TM-20 construct and verifying the presence of the PDGFβR-TM-20 gene, all tags and sites, the plasmid was transformed into BL21(DE3)PLysS cells as previously described (section 2.5.3). The protein was then expressed in minimal media using the conditions optimised for the pET30a-PDGFβR-TM-20 construct. A band was observed close to the predicted molecular weight of 10.85 kDa on the gel (Figure 4.14a) and anti-His western blot (Figure 4.14b) indicating that the protein was expressed (bands highlighted in red box).

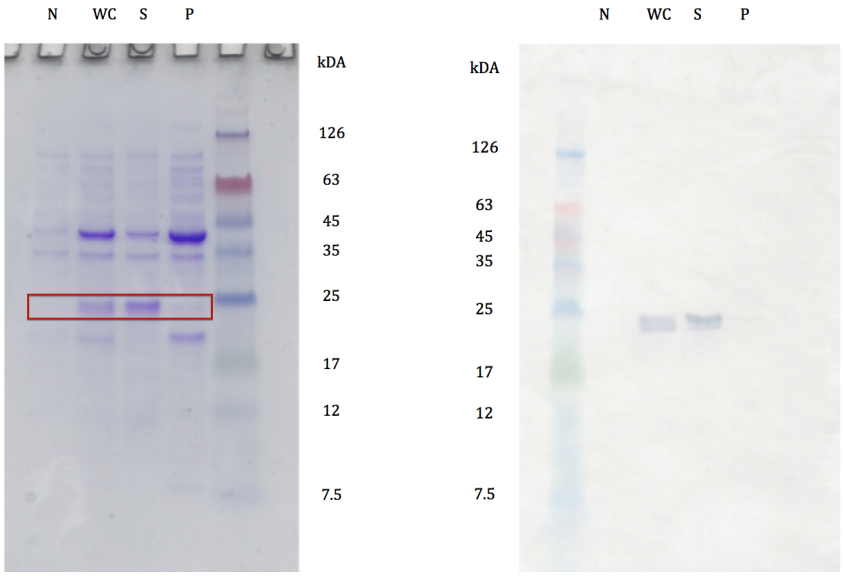


(a) SDS-PAGE **(b)** Western blot

Figure 4.14: Expression check for pET151-PDGFβR-TM-20. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Cells were grown at 37 °C for 4 hours after induction with 1 mM IPTG. Red box highlights the position of the expected monomer PDGFβR-TM-20 band.WC: whole cells, P: pellet after lysis and centrifugation of whole cells, S: supernatant after lysis and centrifugation of whole cells, M: protein ladder.

TEV protease

The PDGFβR-TM-20 protein produced after expression contains a His-tag which is cleavable at a TEV cleavage site. The TEV protease genetic construct for the in-house production of TEV protease was present in a pRK793 plasmid which had been transformed into *E. coli* BL21(DE3)-RIL cells and was expressed as detailed in section 2.7.1. Figures 4.15a & 4.15b show the gel and the anti-His Western blots of the expression check, respectively. It can be seen (bands highlighted in red box) that good expression levels are obtained and all the TEV protease expressed is in the soluble fraction.



(a) SDS-PAGE **(b)** Western blot

Figure 4.15: Expression check for TEV protease. Cells were grown at 15 °C overnight after induction with 0.11 mM IPTG. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, S: supernatant after lysis and centrifugation of whole cells, M: protein ladder.

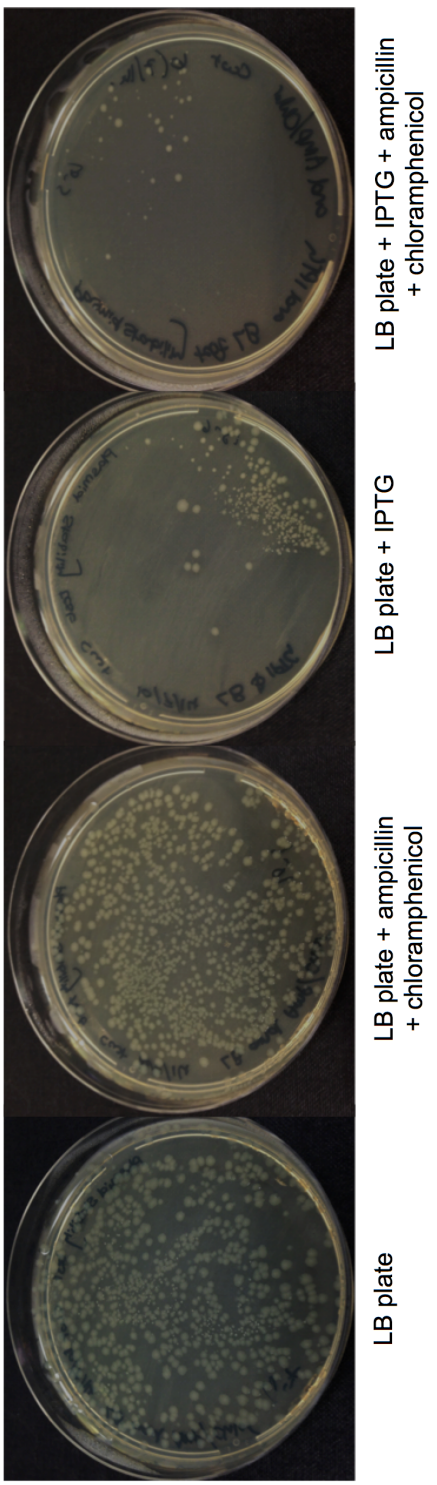
4.5 Plasmid Stability Test

Unlike pET30a which is kanamycin resistant, the pET151 vector is ampicillin resistant and as such requires that plasmid maintenance be monitored closely. This is because the resistance protein, β-lactamase, is secreted which could result in the complete degradation of the of ampicillin present in the bacterial culture (Robinson, 2011). Plasmid maintenance during outgrowths was examined using the plasmid stability test, outlined in section 2.6.5. The cells were plated as outlined in the protocol on four different plates (Table 4.1).

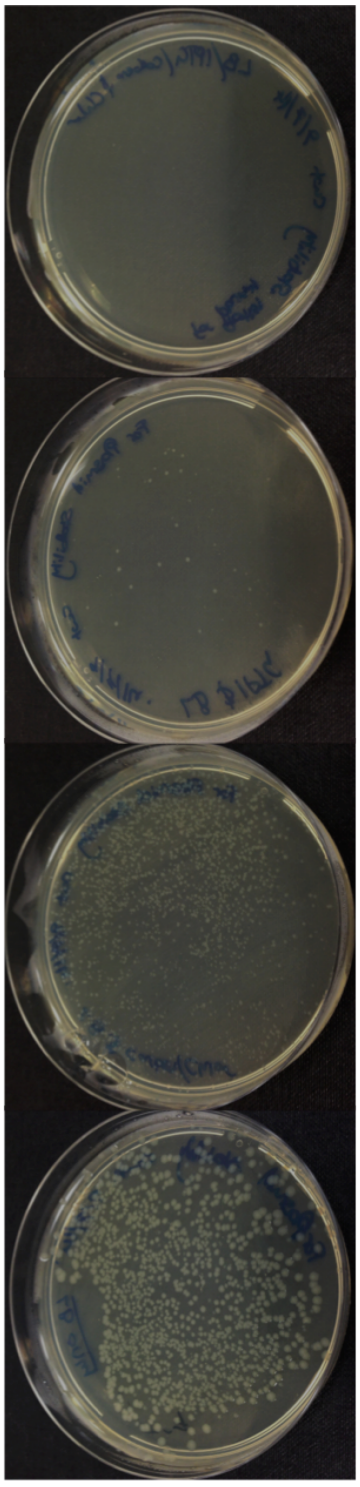
Plate	Cells that grow
LB	All viable cells
LB + antibiotic	Cells that still carry the plasmid
LB + IPTG (1 mM)	Cells that have lost the plasmid or mutants that have lost the ability to express the target gene
LB + antibiotic + IPTG (1 mM)	Mutants that retain the plasmid but have lost the ability to express the target gene

Table 4.1: Plasmid stability test

Following cell outgrowths using the existing protocol, the plasmid stability test was performed and it was observed that a large number of colonies grew in the presence of IPTG (Figure 4.16a). This suggests the cells have lost the protein production plasmid, because in the presence of IPTG, cells dedicate all their resources to protein production rather than cell maintenance. Approximately 45 colonies were also observed in the presence of IPTG as well as antibiotics. While plasmids are retained in these colonies, they have lost their ability to express the protein.



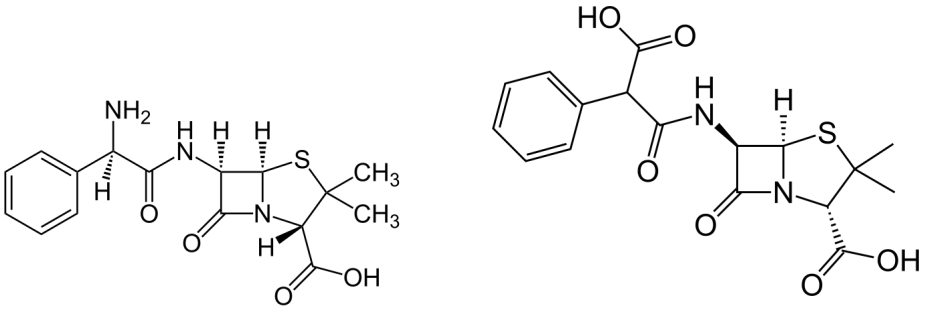
(a) Ampicillin



(b) Carbenecillin

Figure 4.16: Plasmid stability test

In order to minimise plasmid loss, and hence maximise protein yield, there were three main changes that were made to the pET151-PDGFβR-TM-20 expression protocol. First, carbenecillin was substituted for ampicillin. Both ampicillin and carbenicillin form a part of the penicillin family and hence adopt the same mode of action in inhibiting cell wall synthesis. Carbenicillin is a carboxypenicllin whereas ampicillin is an aminopenicillin (Figures 4.17a & 4.17b). Carbenicillin is known to have improved stability over ampicillin when used in growth media. It demonstrates greater resistance to heat and exhibits less degradation at low pH over time making it particularly useful for large-scale liquid culture growth.



(a) Ampicillin structure **(b)** Carbenicillin structure

Figure 4.17: Structures of ampicillin and carbenicillin.

Second, β-lactamase was removed from overnight inoculants. β-lactamase, which is expressed from the *bla* gene in the plasmid, hydrolyses and inactivates ampicillin. It is secreted by the bacteria and hence is present in the culture medium. If overnight inoculants are directly added to the bacterial culture, a certain proportion of ampicillin is immediately deactivated from the β-lactamase present in the overnight inoculant. In order to avoid this, overnight inoculants were gently centrifuged and the bacterial cells were resuspended in fresh medium.

Finally, cells were harvested/induced at a lower OD_{600}. It was also found that at an $OD_{600} > 0.5$, cells are much more likely to lose their plasmid. Hence, when expressing in minimal media, the protocol was modified to harvest all cells at an OD_{600} between 0.4–0.5 before expressing using the high-cell-density expression method (section 4.3.7). When expressing in rich medium, cells were induced at an OD_{600} of 0.5.

The plasmid stability was performed again after incorporating these changes into the protocol. It can be seen in Figure 4.16b that the number of colonies which grow in the presence of IPTG only is greatly reduced. Furthermore, no colonies are observed in the presence of IPTG as well as antibiotics. These observations confirm that the changes made in the protocol have improved plasmid maintenance.

4.6 Mistic-tag to facilitate protein expression

Mistic was first discovered as a membrane-associated protein in *Bacillus subtilis* (Roosild et al., 2005). When used as a fusion partner, it has been shown to greatly improve expression level of integral membrane proteins by overcoming toxicity issues associated with overloading the protein translation machinery. It does not possess a recognisable signal sequence and it is likely that it by-passes *E. coli* translocon machinery and autonomously associates with the bacterial membrane (Dvir and Choe, 2009).

Figure 4.18: NMR structure of Mistic. Adapted from (Roosild et al., 2005).

It was thought that in this study, this property of Mistic could be exploited to express PDGFβR-TM-20 in the bacterial membrane. Additionally, the enhancement of expression that this may provide could allow sufficient quantities to be obtained for NMR studies. This would not only simplify the purification procedure, but would also allow the study of the structure and dynamics of the PDGFβR-TM-20 dimers in native bacterial membranes using solid-state NMR.

4.6.1 Amplification of the PDGFβR-TM-20 gene

Firstly, the PDGFβR-TM-20 gene was amplified from the pET30a-PDGFβR-TM-20 plasmid using the forward and reverse oligonucleotides shown in section 2.3.1. These were designed against the template PDGFβR-TM-20 sequence with the forward primer incorporating *PspXI* restriction site and the reverse primer incorporating *SalI* restriction site. This allowed amplification of the PDGFβR-TM-20 gene by PCR (Figure 4.19) with restriction sites at either end. A TEV protease cleavage site was also incorporated in the reverse primer so as to cleave the His and Mistic tags. These were then double digested with *PspXI* and *SalI* restriction endonucleases to provide sticky ends for ligation.

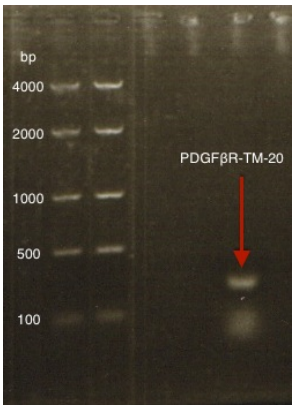


Figure 4.19: Amplification of PDGFβR-TM-20 gene from the pET30a-PDGFβR-TM-20 plasmid.

4.6.2 Ligation

The Mistic gene was contained in a Gateway cloning pDEST17 vector provided by Florent Bernaudat (Institut de Biologie Structurale Jean-Pierre Ebel, CEA, Grenoble, France) as used previously (Bernaudat et al., 2011a) . The plasmid was transformed into One Shot® *ccd*B Survival^{TM}2 T1^{R} Chemically Competent Cells from Life Technologies, UK. They were then propagated in a rich medium and the plasmid was extracted and sent for sequencing to ensure plasmid integrity.

To generate compatible sticky ends for ligation, the plasmid was double digested with *PspXI* and *SalI* restriction endonucleases. Figure 4.20 shows an agarose gel

with bands close to 5114 bp and 1570 bp representing the cut plasmid and cut fragment respectively.

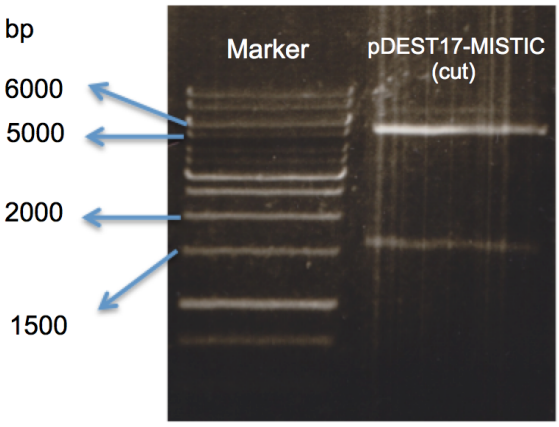


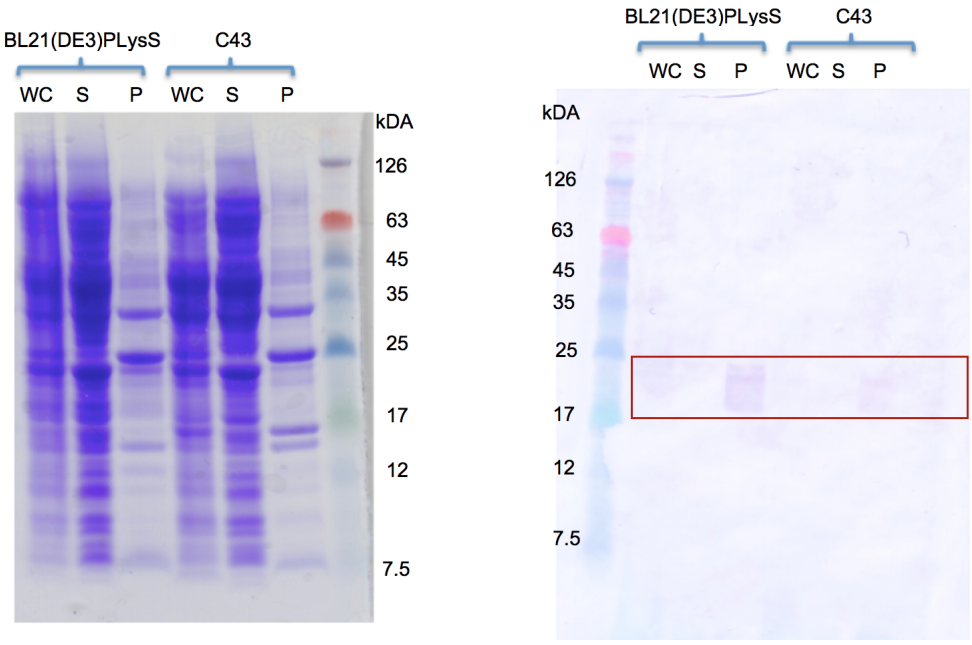
Figure 4.20: The pDEST17-Mistic plasmid double digested at the *SalI* and *PspXI* restriction sites. Bands are seen close to expected values of 5114 bp and 1570 pb representing the cut plasmid and cut fragment respectively.

The double digested PCR products and the pDEST17-Mistic plasmid were then mixed and ligated as per the protocol detailed in section 2.3.7. The mixture was then used to transform *E. coli* TOP10 cells (section 2.5.3). Positive transformants were grown to extract DNA, which was sent for sequencing to verify the presence of the insert and to ensure no mutations were present.

4.6.3 Expression of pDEST17-Mistic construct

The pDEST17-Mistic-PDGFβR-TM-20 construct was transformed into C43 and BL21(DE3)PLysS cells to check for expression levels and to also ensure that the protein was being inserted into the membrane due to the Mistic tag. Following transformation, Mistic-PDGFβR-TM-20 was expressed in rich medium in the first instance (as per existing protocol). Figures 4.21a & 4.21b show the SDS-PAGE gel and anti-His western blot respectively for the expression check. Even though both cells expressed the fusion construct, disappointingly, none of it was found in the soluble fraction following lysis which contains the bacterial membranes; all of it had gone into inclusion bodies. However, the expression level in BL21(DE3)PLysS cells

showed significantly higher expression levels than C43 cell line.



(a) SDS-PAGE **(b)** Western Blot

Figure 4.21: Expression check for pDEST17-Mistic-PDGFβR-TM-20. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Cells were grown at 37 ° for 4 hours after induction with 1 mM IPTG. Red box highlights the position of the expected monomer band. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, S: supernatant after lysis and centrifugation of whole cells.

This test expression was performed at 37 °C. It has been observed for both BL21(DE3) and C43 cell lines that solubility of membrane proteins and expression level can be greatly improved by expressing the protein at lower temperatures (Berthold et al., 2011). In light of this, further expression trials were performed at 30 °C and 15 °C. Figure 4.22 shows the western blot obtained following expression. Cells were induced for 4 hours at 37 °C, 8 hours at 30 °C and 24 hours at 15 °C. Even though overall expression levels were improved by lowering the temperatures to 30 °C, no protein was observed in the fraction containing the bacterial membranes. No expression was observed at 15 °C.

It is unclear why the protein was being directed to inclusion bodies despite the Mistic tag having an overall hydrophilic composition. It could be that Mistic was not folding properly when fused with the peptide, or it could be that PDGFβR-TM-20 is unable

to bypass the *E. coli* translocon machinery when fused with Mistic. Due to time constraints, further optimisations to troubleshoot this could not be performed.

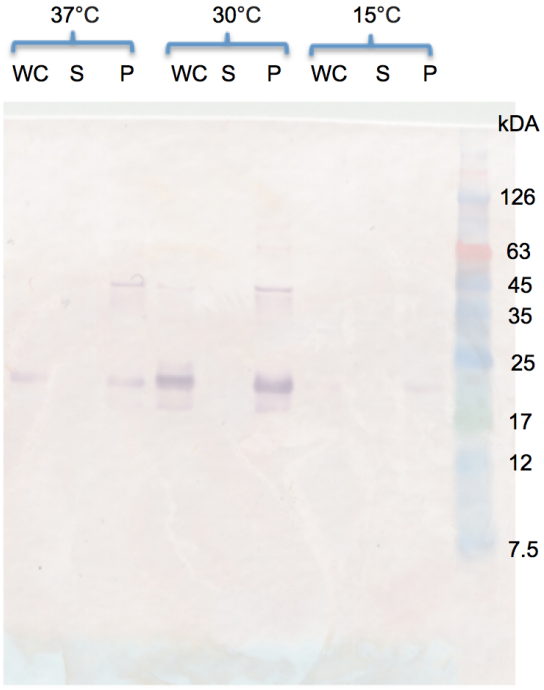


Figure 4.22: Expression of pDEST17-Mistic-PDGFβR-TM-20 at various temperatures observed on an anti-His western blot. WC: whole cells, P: pellet after lysis and centrifugation of whole cells, S: supernatant after lysis and centrifugation of whole cells.

4.7 Summary and Conclusion

This chapter focussed on methods adopted and optimisations performed to maximise the yield of PDGFβR-TM-20 such that a functionally relevant sample of PDGFβR-TM-20 could eventually be prepared for solution- and solid-state NMR spectroscopy. A plasmid construct pET30a-PDGFβR-TM-20 which had previously been used in the Dixon lab to produce PDGFβR-TM-20 for biophysical analyses was chosen as a starting point for this purpose.

A number of cell lines were tested and it was determined that the best cell line for expressing PDGFβR-TM-20 was BL21(DE3)PLysS. The expression was subsequently optimised by fine-tuning the glucose concentration, temperature and IPTG

concentration. The effect on protein expression due to the addition of BME vitamins and iron (III) chloride was also assessed. Furthermore, to maximise protein yield, high-cell-density method of expression was adopted and optimised.

However, due to problems in cleaving the His-tag from the pET30a-PDGFβR-TM-20 construct (section 5.2.4), a related pET vector which utilises the TOPO cloning method for entry clone insertion was chosen to take the study forward. This method simplifies and improves the cloning process by using a cloning vector that has been "activated" with topoisomerase I in lieu of ligase and long ligation times. Since the TOPO cloning vector chosen, pET151, is also a pET vector, it was hoped that similar level of expression would be obtained using the optimisations that were performed on the pET30a vector. This was indeed the case; however, since pET151 carries the gene for AmpR in place of KanR, it was necessary to ensure plasmid stability as ampicillin is much more susceptible to degradation than kanamycin. This was done by doing plasmid stability tests and modifying the expression protocol to ensure plasmid maintenance. Since the pET carries the TEV protease cleavage site, it was also ensured that sufficient quantities of TEV protease can be produced using the in-house construct.

Another avenue which was explored to express PDGFβR-TM-20 in bacterial membranes was fuse it with a Mistic-tag which has been known to by-pass *E. coli* translocon machinery and directly associate with the inner bacterial membrane. This would simplify the purification procedure and also grant the added potential of being able to look at the structural properties of PDGFβR-TM-20 in native membranes using solid-state NMR. Traditional cloning procedures were followed to create a genetic construct which expresses PDGFβR-TM-20 fused with Mistic. However, the fusion construct only expressed as inclusion bodies despite some optimisations to increase the solubility of the protein.

Chapter 5

PDGFβR-TM-20 Purification, Sample Preparation and Characterisation

5.1 Introduction

In the previous chapter, protocols and strategies were outlined that were optimised to obtain the maximum yields of PDGFβR-TM-20 for NMR analyses. This chapter will begin by detailing the methods adopted for the purification of the PDGFβR-TM-20. Particularly, there will be a focus on the problems encountered in cleaving the His-tag from the protein and how they were overcome. Furthermore, the protocols devised for the reconstitution of PDGFβR-TM-20 into membrane mimetics for solution- and solid-state NMR will be presented along with some preliminary biophysical characterisation of the samples.

As explained in the previous chapter, there were two plasmid constructs used to express PDGFβR-TM-20: pET30a-PDGFβR-TM-20 and pET151-PDGFβR-TM-20. The former produced a protein which contained a proteolytic cleavage site for recombinant enterokinase (rEK) and the latter contained a TEV protease cleavage site. These will henceforth be referred to as rEK-PDGFβR-TM-20 and TEV-PDGFβR-TM-20 respectively.

5.2 Purification and Cleavage of rEK-PDGFβR-TM-20

5.2.1 The purification strategy

The protocol used previously in the lab for preparing PDGFβR-TM-20 in detergent micelles from inclusion bodies involved a series of purification steps to obtain PDGFβR-TM-20–His_6. The expression region in pET30a-PDGFβR-TM-20 encodes for the following sequence:

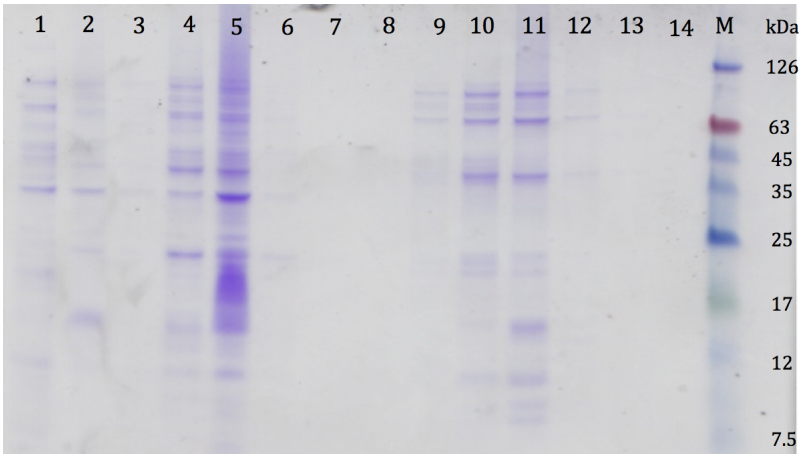
$$His_6\text{–}S\text{-}tag\text{–}thrombin\text{–}enterokinase\text{–}PDGF\beta R\text{-}TM\text{-}20\text{–}His_6.$$

The purification protocol involved solubilising the rEK-PDGFβR-TM-20 inclusion bodies in Guanidine HCl and then performing IMAC followed by on-column refolding of the protein using *n*-Dodecyl β-D-maltoside (DDM). The purified protein was then cleaved using recombinant enterokinase yielding His_6–S-tag–thrombin–enterokinase and PDGFβR-TM-20–His_6. Pure PDGFβR-TM-20–His_6 was then obtained by passing the mixture through an S-tag column.

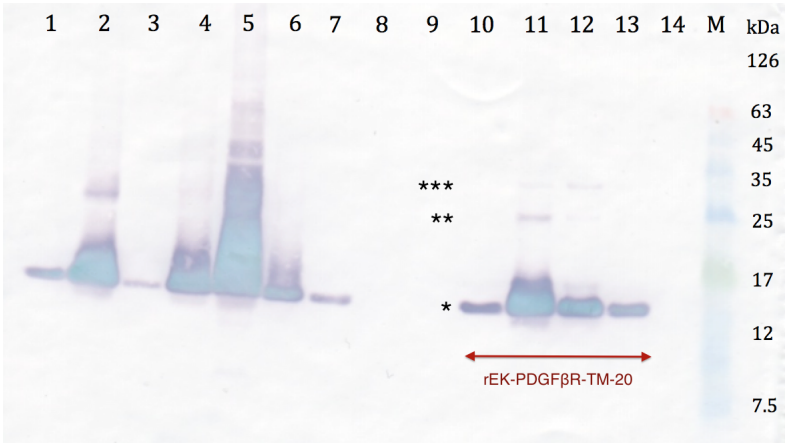
5.2.2 Purification of rEK-PDGFβR-TM-20

The protein was first expressed in LB and trial purifications were carried out as per the existing protocol with two exceptions. Firstly, the choice of detergent was changed to *n*-Decyl β-D-maltoside (DM) as it has a ten-fold higher critical micelle concentration (CMC) and hence could be more easily removed through dialysis for reconstitution into lipid bilayers. Secondly, the denaturing agent was changed to urea in lieu of Guanidine HCl as the latter precipitates in the presence of SDS and leads to various artefacts on an SDS-PAGE gel making the monitoring of expensive labelled proteins in solution difficult.

The protocol has been outlined in detail in section 2.6.2. Aliquots were taken at each stage of the purification and run on an SDS-PAGE gel followed by a western blot (Figures 5.1a & 5.1b respectively). Several observations were made during the trial purifications. First, it was seen (Lane 5, Figure 5.1b) that upon leaving the



(a) SDS-PAGE



(b) Western Blot

Figure 5.1: Purification of rEK-PDGFβR-TM-20. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Lane 1: supernatant after lysis and centrifugation, Lane 2: pellet after lysis and centrifugation, Lane 3: supernatant after centrifuging again, Lane 4: supernatant after solubilisation of inclusion bodies in urea and centrifugation, Lane 5: pellet after solubilisation of inclusion bodies in urea and centrifugation, Lane 6: 1st flow-through with DM in buffer , Lane 7: 2nd flow-through with DM in buffer, 8 = 3rd flow-through with DM in buffer + 10 mM imidazole, Lane 9–12: flow-throughs with DM in buffer + 300 mM imidazole, Lane 13: MES wash, Lane 14: dH_2O wash. The positions of monomer, dimer and trimer are marked with *, ** and *** respectively.

inclusion bodies in urea for ∼2 hours as per protocol for GnHCl, not all of them were solubilised. It was observed in subsequent purifications that leaving the inclusion bodies in urea solution in an end-over-end mixer overnight completely solubilised them. Using urea instead of GnHCl was particularly useful in ensuring that all of the labelled proteins had been extracted through IMAC purification when employing batch binding.

After purification through IMAC, several bands were observed on the SDS-PAGE gel which do not correspond to the protein of interest and are indicative of non-specific binding to the resin (Lanes 9–12, Figure 5.1a) . This was significantly reduced by making three changes. Firstly, 10 mM imidazole was added to the solution during binding to the resin. Secondly, the imidazole concentration was increased during the wash stage from 10 mM to 25 mM. Thirdly, the wash volume was increased from 1 column volume to 10 column volumes. All of these changes showed no deleterious effects on protein yield.

The volume of elution buffer was increased due to residual protein bound to the resin after passing three column volumes of elution buffer through it (Lane 13, Figure 5.1b). Six column volumes ensured complete elution of the protein.

Bio-Beads could also be employed to remove detergent as the protein is reconstituted into liposomes. The ideal detergent for this purpose is Triton X-100 and purification was also carried out using Triton X-100 without any loss of protein yield. Indeed, the protein exhibited greater long term stability in Triton X-100.

5.2.3 Protein quantification using BCA assay

Concentrations of both PDGFβR-TM-20 and the TEV protease were calculated using the BCA assay as detailed in section 2.8.6. Protein concentration was calculated by determining the gradient of the standard curve (m) and applying the formula:

$$\text{Concentration } x = \frac{A_{562}}{m}$$

where A_{562} is the absorbance at 562 nm. Figure 5.2 shows a typical BCA assay standard curve.

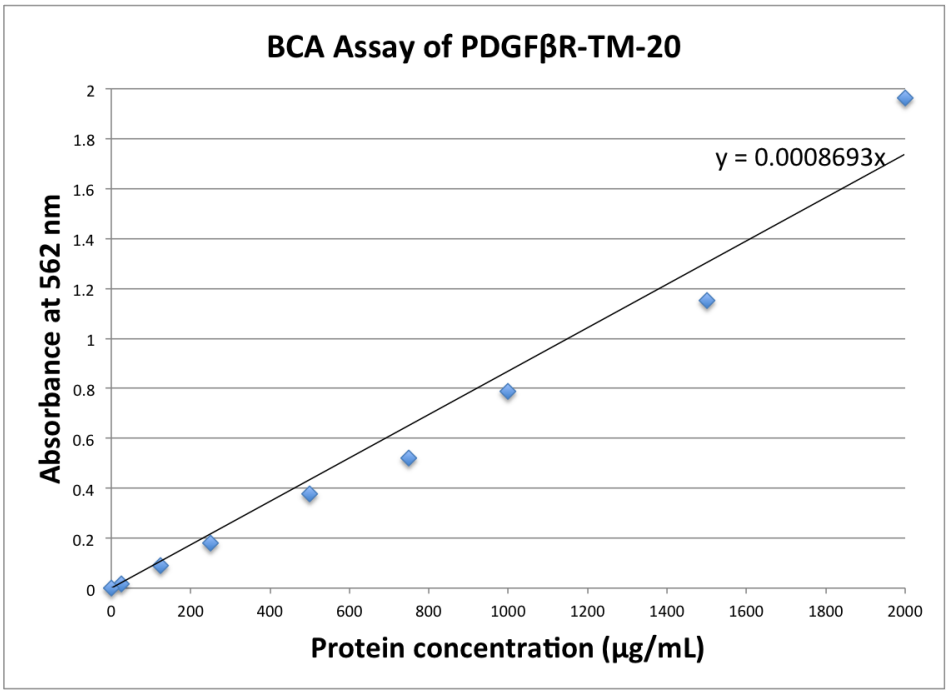


Figure 5.2: A typical BCA assay standard curve obtained by measuring the known concentrations of BSA

5.2.4 Site-specific enzymatic cleavage of rEK-PDGFβR-TM-20

As mentioned earlier, the pET30a plasmid expressed PDGFβR-TM-20 with tags cleavable by recombinant enterokinase (rEK). Initially, an enterokinase cleavage capture kit was used to test for cleavage of rEK-PDGFβR-TM-20 in Triton X-100 detergent. The cleavage of 20 μg, 50 μg and 100 μg of the protein was tested with a single unit of enterokinase as per the manufacturer's instructions. Aliquots were taken out at 2, 4, 8 and 16 hours after starting the reaction.

An SDS-PAGE gel was run for each of the aliquots and is shown in Figure 5.3. The theoretical masses for the uncleaved protein, cleaved protein and cleaved tag are 13.86 kDa, 9.46 kDa and 4.40 kDa respectively. Complete cleavage was only observed after 16 hours for 20 μg and 50 μg of protein. This would suggest that the entire cleavage kit (50 units) can only cleave 2.5 mg of protein. To ensure that the low cleavage efficiency was not due to the degradation of the protease, a new kit was

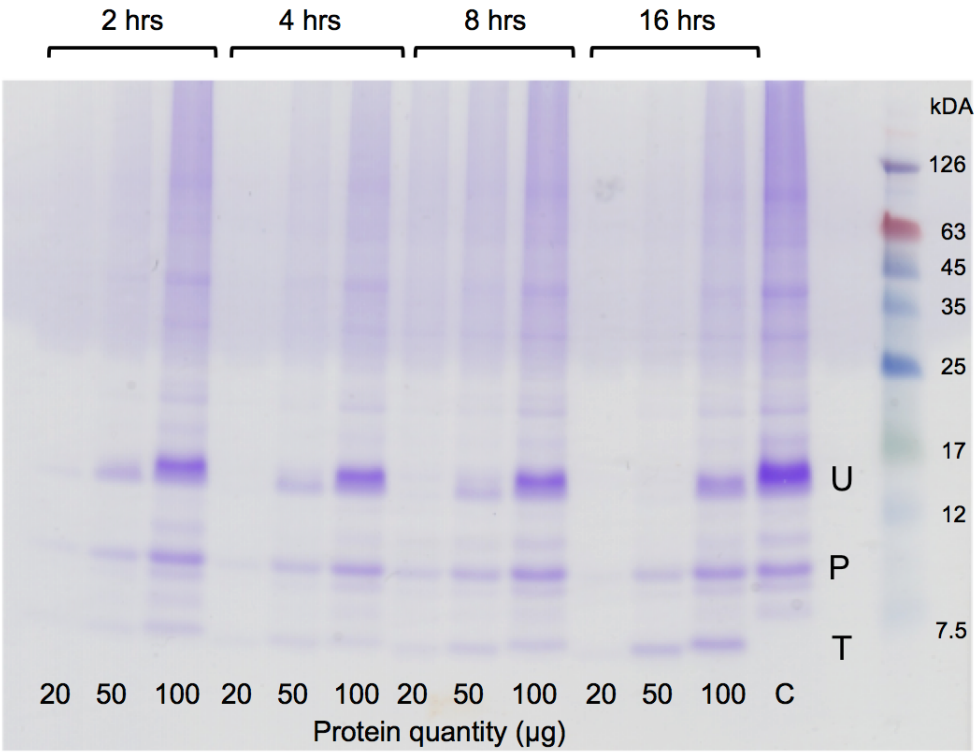
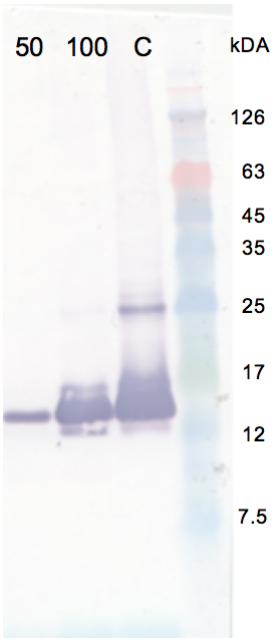


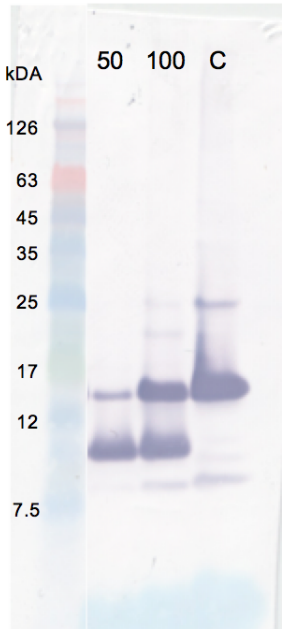
Figure 5.3: rEK-PDGFβR-TM-20 cleavage using enterokinase after 2, 4, 8 and 16 hours for 20 μg, 50 μg and 100 μg of the protein. U: Uncleaved rEK-PDGFβR-TM-20, P: cleaved PDGFβR-TM-20, T: cleaved tag. Lane C represents the control with no protease.

also tested. Figure 5.4a shows the anti-His western blot obtained after performing the cleavage reaction on 50 μg and 100 μg of the protein with 1 unit of enterokinase for 16 hours. Even though the cleaved protein is His-tagged, the cleaved protein could not be observed on the blot. The same blot was developed with C-terminal specific anti-His antibody and the cleaved protein could be observed on the blot (Figure 5.4b). However, no improvement in cleavage efficiency was observed; 100 μg of the protein could still not be cleaved with 1 unit of the enterokinase. Such low cleavage efficiency renders this protocol of producing large quantities of PDGFβR-TM-20 for NMR studies infeasible.

In light of this, a new plasmid with a TEV cleavage site was introduced. This choice was based on the fact that TEV was currently being produced in the lab in large quantities in an economical manner. This also allowed the non-cleavable C-terminal His-tag to be removed for simplicity of purification.



(a) rEK cleavage using anti-His antibody



(b) rEK cleavage using C-terminal specific anti-His antibody

Figure 5.4: rEK-PDGFβR-TM-20 cleavage using enterokinase after 16 hours at room temperature for 50 μg and 100 μg of the protein. Lane C represents the control with no protease.

5.3 Purification and Cleavage of TEV-PDGFβR-TM-20

5.3.1 The purification strategy

The protocol used for the purification of TEV-PDGFβR-TM-20 was exactly the same as that used for rEK-PDGFβR-TM-20, detailed in section 5.2.2. The expression region in pET151-PDGFβR-TM-20 encodes for the following sequence:

$$His_6\text{–}V5\text{-}tag\text{–}TEV\text{–}PDGF\beta R\text{-}TM\text{-}20.$$

The strategy was to add TEV protease to the purified protein in a 1 : 1 ratio and, after cleavage, remove the cleaved tag plus the His-tagged TEV protease using IMAC.

5.3.2 Purification of TEV-PDGFβR-TM-20

Purification after expression of TEV-PDGFβR-TM-20 was performed as described in section 2.6.2. Aliquots were collected at each stage of the purification and run on an SDS-PAGE gel (Figure 5.5). A clear thick band can be seen in the elution lanes close to the expected molecular weight of 10.85 kDa.

5.3.3 Purification of TEV protease

TEV protease was purified as detailed in section 2.7.2. Aliquots were taken from the washing steps as well as the elution steps and run on an SDS-PAGE gel and western blotted (Figures 5.6a & 5.6b). Pure TEV indicated by an (*) can be observed in the elution lanes 2 & 3 along with a dimer species indicated by an (**). This construct of TEV employs a maltose-binding-protein (MBP) fusion partner which self-cleaves post-expression. The monomer and dimer bands for this can also be observed, indicated with (+) and (++) respectively.

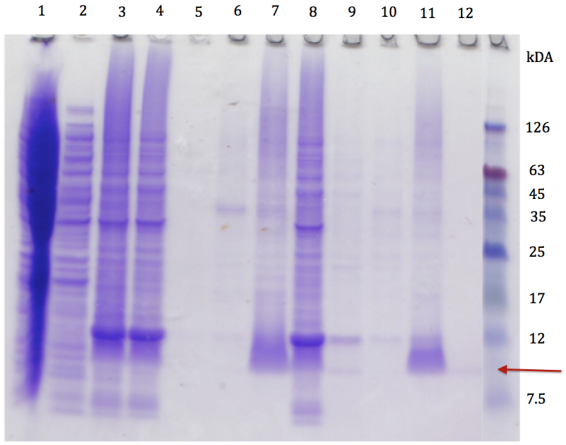
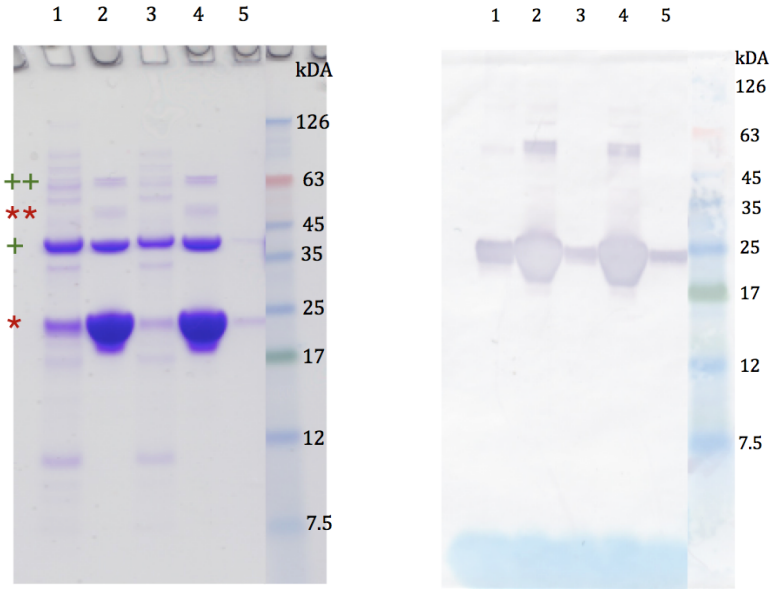


Figure 5.5: Purification of TEV-PDGFβR-TM-20. Lane 1: supernatant after lysis and centrifugation, Lane 2: supernatant after inclusion body wash, Lane 3: after filter sterilisation, Lane 4: protein remaining in solution after first batch-binding, Lane 5: wash and on-column refolding using Triton X-100, Lane 6: 25 mM imidazole wash, Lane 7: elution using 300 mM imidazole, Lane 8: protein remaining in solution after first second batch binding step, Lane 9: wash and on-column refolding using Triton X-100, Lane 10: 25 mM imidazole wash, Lane 11: elution using 300 mM imidazole, Lane 12: MES wash. Position of the TEV-PDGFβR-TM-20 band is marked with a red arrow.



(a) SDS-PAGE **(b)** Western Blot

Figure 5.6: Purification of TEV. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Lane 1: wash elution (1st batch), Lane 2: Protein elution (1st batch), Lane 3: wash elution (2nd batch), Lane 4: protein elution (2nd batch). Lane 5: MES wash. TEV protease monomer and dimer indicated by * and ** respectively. Cleaved MBP fusion partner monomer and dimer indicated by + and ++ respectively.

5.3.4 Site-specific enzymatic tag-cleavage of TEV-PDGFβR-TM-20 using MBP-TEV

After both TEV-PDGFβR-TM-20 and TEV protease were purified, they were mixed together in a 1 : 1 protein to protease molar ratio to perform the cleavage reaction at room temperature overnight. Aliquots were collected and run on an SDS-PAGE gel (Figure 5.7). It was observed that after leaving the mixture overnight at room temperature, cleavage of the tag had occurred. The TEV-PDGFβR-TM-20 band close to the expected molecular weight value of 10.85 kDa (indicated with U in Figure 5.7) had largely disappeared and a new band near the theoretical molecular weight value of 7.83 kDa for the cleaved protein appeared (indicated with C in Figure 5.7). A second IMAC step was performed to remove the tags and obtain pure PDGFβR-TM-20. The TEV protease and the cleaved tag from TEV-PDGFβR-TM-20 were removed completely. However, the cleaved MBP fusion partner was not removed after the second IMAC step.

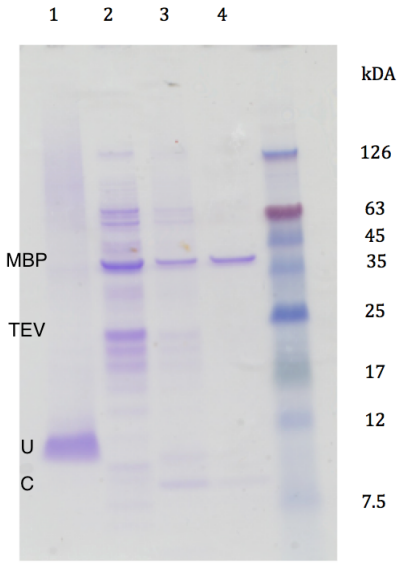


Figure 5.7: Proteolytic cleavage of TEV-PDGFβR-TM-20. Lane 1: purified uncleaved TEV-PDGFβR-TM-20. Lane 2: TEV protease solution, Lane 3: post-cleavage reaction, Lane 4: After final IMAC. U: uncleaved TEV-PDGFβR-TM-20, C: cleaved TEV-PDGFβR-TM-20

Separation of TEV protease from MBP using gel-filtration chromatography

Since MBP could not be removed from the cleavage reaction, an attempt was made to remove the MBP from the TEV protease purified solution using gel-filtraton chromatography as this method had been employed previously by others (Tropea et al., 2009). Absorbance at 280 nm and 220 nm was monitored in the chromatogram, and aliquots were taken from the fractions which were thought to contain the TEV protein. These were then run on an SDS-PAGE gel and can be seen in Figure 5.8. Unfortunately, MBP was co-eluting with TEV protease throughout the separation and could not be separated.

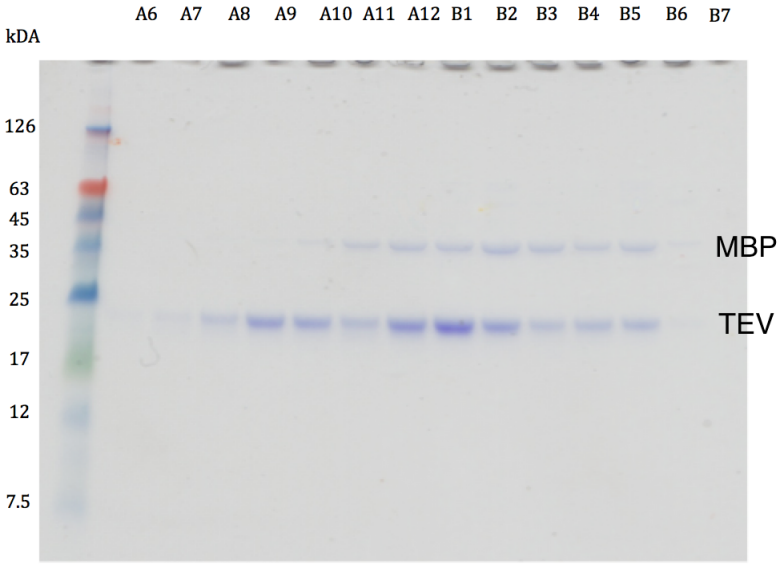


Figure 5.8: Gel-filtration profile of the TEV/MBP mixture

Cleavage attempts using TEV protease from other sources

In light of the above, it was decided to obtain TEV protease from another source and test the cleavage efficiency, in the first instance. One batch of TEV protease was obtained from Dr David Roper's group at the University of Warwick and the other was a commercially available TEV protease purchased from Sigma Aldrich, UK. Cleavage reactions were initiated with 1 : 1, 1 : 2, 1 : 5 and 1 : 10 TEV

protease to TEV-PDGFβR-TM-20 ratios. Aliquots were taken after 16 hours and run on an SDS-PAGE gel (Figure 5.9). Unfortunately, the commercially available TEV protease did not cleave at all. The TEV protease from Dr David Roper's only showed signs of cleavage when added in a 1 : 1 ratio.

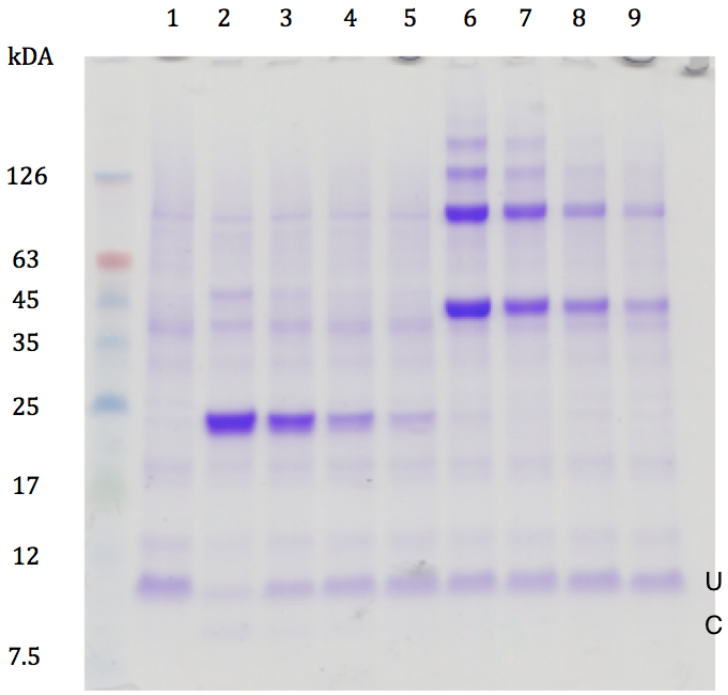


Figure 5.9: Proteolytic cleavage of TEV-PDGFβR-TM-20. Lane 1: purified uncleaved TEV-PDGFβR-TM-20. Lanes 2–5: TEV from Dr David Roper's lab with TEV protease to TEV-PDGFβR-TM-20 ratios of 1 : 1, 1 : 2, 1 : 5 and 1 : 10, Labes 6–9: TEV from Sigma Aldrich with TEV protease to TEV-PDGFβR-TM-20 ratios of 1 : 1, 1 : 2, 1 : 5 and 1 : 10. U: uncleaved TEV-PDGFβR-TM-20, C: cleaved TEV-PDGFβR-TM-20

Purification of MBP/TEV mixture using amylose resin

MBP has been used as a fusion protein in expressing proteins for decades and purification of these proteins has often exploited the MBP's affinity for amylose (di Guana et al., 1988; Maina et al., 1988). It was proposed that MBP-TEV mixture could be passed through an amylose column allowing the MBP to preferentially bind to the amylose resin and hence removing it from the solution. Aliquots were collected and an SDS-PAGE gel was run (Figure 5.10). Unfortunately, not only did all of the

MBP bind to the amylose resin, but all of the TEV also bound irreversibly to the resin and hence could not be separated from the MBP using this method.

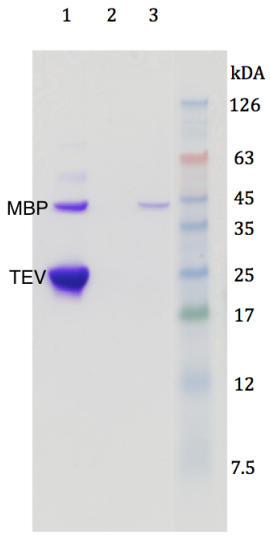
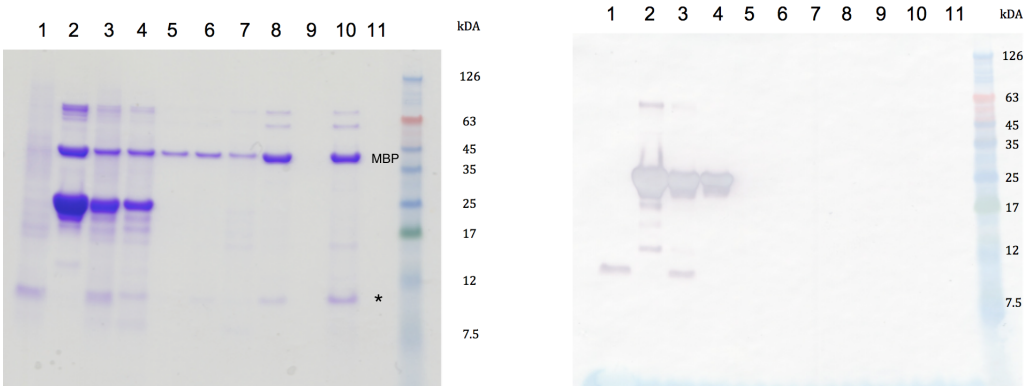


Figure 5.10: TEV protease purification from MBP using amylose resin. Lane 1: untreated TEV/MBP mixture, Lane 2: after passing through amylose resin. Lane 3: elution with 50 mM maltose.

Removal of MBP using centrifugal concentrators

It was also proposed that the large difference in molecular weight could be exploited to separate PDGFβR-TM-20 from MBP following the cleavage reaction. This could be performed using centrifugal concentrators with a molecular weight cut-off of 50 kDa, bearing in mind that MBP is ∼42.5 kDa and the PDGFβR-TM-20/Triton X-100 micelle is $>$ 90 kDa. Two separate concentrators were used, one with a PES membrane and the other with a cellulose membrane. Aliquots were collected before and after conecntration and run on an SDS-PAGE gel followed by a western blot as shown in Figures 5.11a & 5.11b. It was observed that separation did not occur using centrifugal concentrators with either of the membranes (Lanes 8 and 10). Furthermore, after centrifugation, at higher concentrations, the cleaved protein ran through the gel a bit more slowly than expected and the protein band (indicated by an asterisk (*)) was observed at a higher molecular weight. No proteins were observed in the filtrate.



(a) SDS-PAGE **(b)** Western Blot

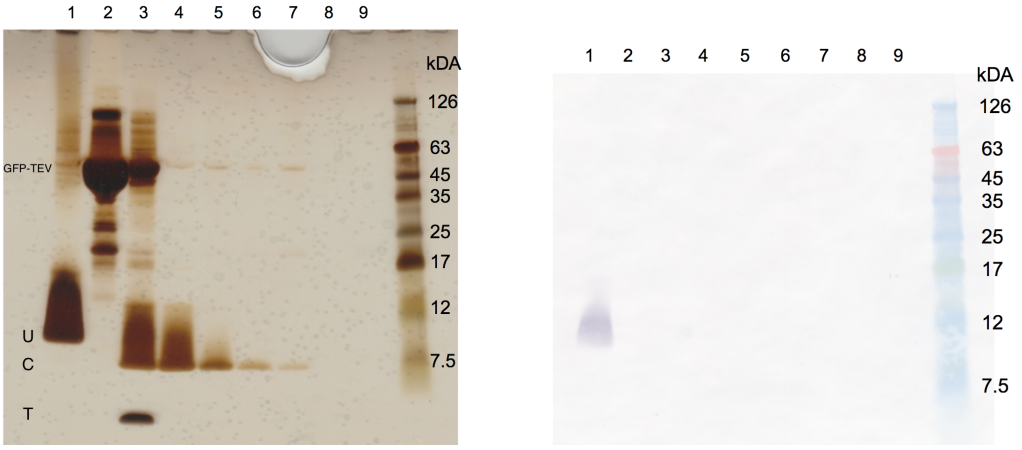
Figure 5.11: TEV protease purification using centrifugal concentrators. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Lane 1: purified uncleaved TEV-PDGFβR-TM-20, Lane 2: purified MBP-TEV, Lane 3: before cleavage, Lane 4: after cleavage, Lane 5: post-cleavage IMAC (first elute), Lane 6: post-cleavage IMAC (second elute), Lane 7: MES wash, Lane 8: after concentrating (PES membrane), Lane 9: filtrate, Lane 10: after concentrating (cellulose membrane), Lane 11: filtrate.

5.3.5 Site-specific enzymatic tag-cleavage of TEV-PDGFβR-TM-20 using GFP-TEV

Since TEV protease exhibits poor solubility when expressed in bacteria, the MBP tag has traditionally been employed to increase the solubility and expression level of TEV protease in *E. coli*. More recently, however, it has been found that even higher levels of TEV protease can be expressed if it is tagged with green fluorescent protein (GFP) (Wu et al., 2010). Additionally, the TEV protease is fully active with the GFP tag bound and there is no need to cleave it from the TEV protease, as opposed to MBP-TEV protease which self-cleaves *in vivo* after expression.

The expression construct for GFP-TEV protease was obtained from Dr Kelley Moremen at the University of Georgia, USA. This construct was expressed at high levels and cleavage was tested. Aliquots were collected at each stage of the cleavage and run on an SDS-PAGE gel (Figure 5.12a). A anti-His western blot (Figure 5.12b) was also performed to monitor cleavage after adding an excess of GFP-TEV to TEV-PDGFβR-TM-20. It can be seen in the western blot that complete cleavage occurred overnight at room temperature (Lane 3). After a second round of IMAC, the cleaved

tag and GFP-TEV were removed to yield purified PDGFβR-TM-20.



(a) SDS-PAGE **(b)** Western Blot

Figure 5.12: TEV-PDGFβR-TM-20 cleavage using GFP-TEV. Panel (a) shows the SDS-PAGE gel whereas B shows the anti-His western blot. Lane 1: purified uncleaved TEV-PDGFβR-TM-20, Lane 2: purified GFP-TEV, Lane 3: after overnight cleavage, Lane 4: after second IMAC, Lane 5: 20 mM Imidazole wash #1, Lane 6: 20 mM Imidazole wash #2, Lane 7: 40 mM Imidazole wash, Lane 8: MES wash #1, Lane 9: MES wash #1. U: uncleaved TEV-PDGFβR-TM-20, C: cleaved TEV-PDGFβR-TM-20, T: cleaved tag.

In order to maximise the use of GFP-TEV protease, a range of different protein to protease ratios were tested. An aliquot was taken from each of the reactions and run on an SDS-PAGE gel (Figure 5.13). It was found that at protein to protease ratio of 1 : 10 (Lane 5)was optimal for cleavage.

The tag and the TEV protease were removed from the solution using IMAC. The resulting PDGFβR-TM-20 in Triton X-100 solution was concentrated tenfold through a centrifugal concentrator and run on a gel, as the protein band after cleavage is very broad and difficult to visualise. After centrifugation, five clear bands were observed in positions corresponding to higher order oligomers of PDGFβR-TM-20 (Figure 5.13, Lane 8). To confirm this, the bands in the gel were sent for mass spectrometry analysis as detailed in section 2.8.5. Mass spectrometry analysis revealed that each of the bands represented various oligomeric states of PDGFβR-TM-20 (Figure 5.14). Even though SDS-PAGE shows the presence of higher order oligomers at these high concentration, it has been shown that PDGFβR-TM exists primarily as dimers in detergent micelles as well as lipid vesicles. (Muhle-Goll et al., 2012).

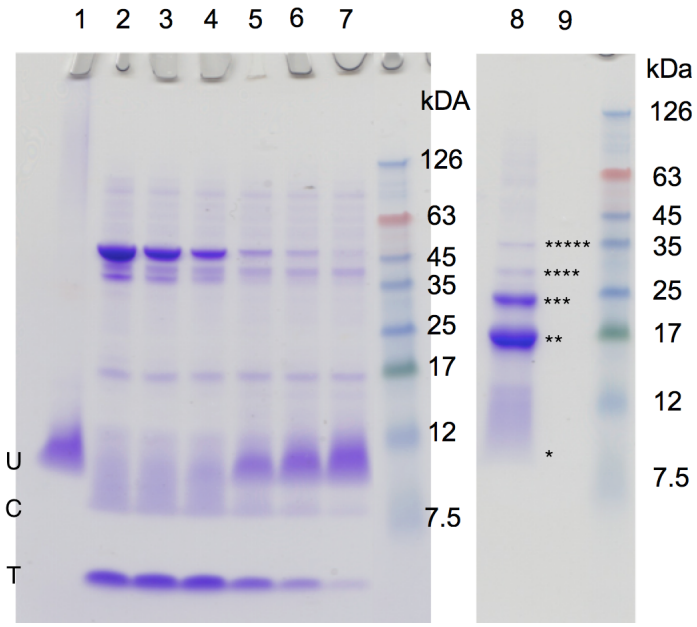


Figure 5.13: GFP-TEV protease cleavage observed on an SDS-PAGE gel to ensure cleavage optimisation. Lane 1: purified uncleaved TEV-PDGFβR-TM-20, Lanes 2–7: GFP-TEV cleavage reactions at a protease to protein ratios of 2 : 1, 1 : 1, 1 : 2, 1 : 10, 1 : 30 and 1 : 60 respectively, Lane 8: concentration of cleaved PDGFβR-TM-20, Lane 9: Filtrate after centrifugal concentration. U: uncleaved TEV-PDGFβR-TM-20, C: cleaved TEV-PDGFβR-TM-20

.

The methods and sample protocols developed here were used to prepare samples for biophysical characterisation.

Figure 5.14: Mass spectrometry analysis of PDGFβR-TM-20 oligomers.

5.4 Sample Preparation for Solid-State NMR

5.4.1 Choice of lipids

It has been observed previously that the PDGFβ receptor is localised to lipid rafts (Pike, 2005) which are cholesterol-rich, highly ordered regions in the membrane with

a high proportion of saturated fatty acyl chains. It has further been hypothesised that the thickness of the membrane controls the ability of the PDGFβR-TM domain to dimerise, and that the activation of the receptor may require its relocation to membrane-thick regions in the form of lipid rafts (Muhle-Goll et al., 2012).

For reconstitution of PDGFβR-TM-20 into liposomes, three different lipids (with different chain lengths and saturation) were used to check how well they would pellet with ultracentrifugation: DMPC (C14:0), POPC (C16:0/C18:1) and DEPC (C22:1 (*cis*)). Knowing this property was important to assess whether they could be centrifuged into a solid-state NMR rotor. Figure 5.15 shows proteoliposome pellets in 1.5 mL ultracentrifuge tubes which have been spun at 267,000 $\times$ g for 45 minutes. DMPC and POPC liposomes formed dense pellets while DEPC liposomes did not pellet. DEPC lipids have the longest chains out of the three lipids, and hence would be useful in mimicking thick lipid rafts, but the PDGFβR-TM in POPC lipids has also been shown to have a tilt angle of 20 $^\circ$ consistent with the dimeric species obtained in DPC micelles (Muhle-Goll et al., 2012). Hence POPC lipids were used for reconstituting PDGFβR-TM-20 into liposomes for CD and NMR analyses.

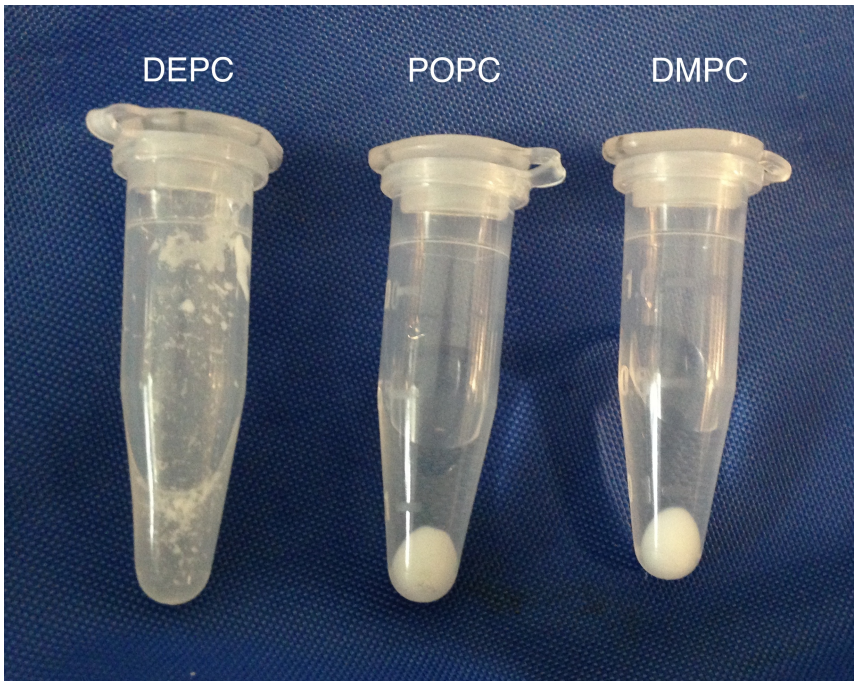
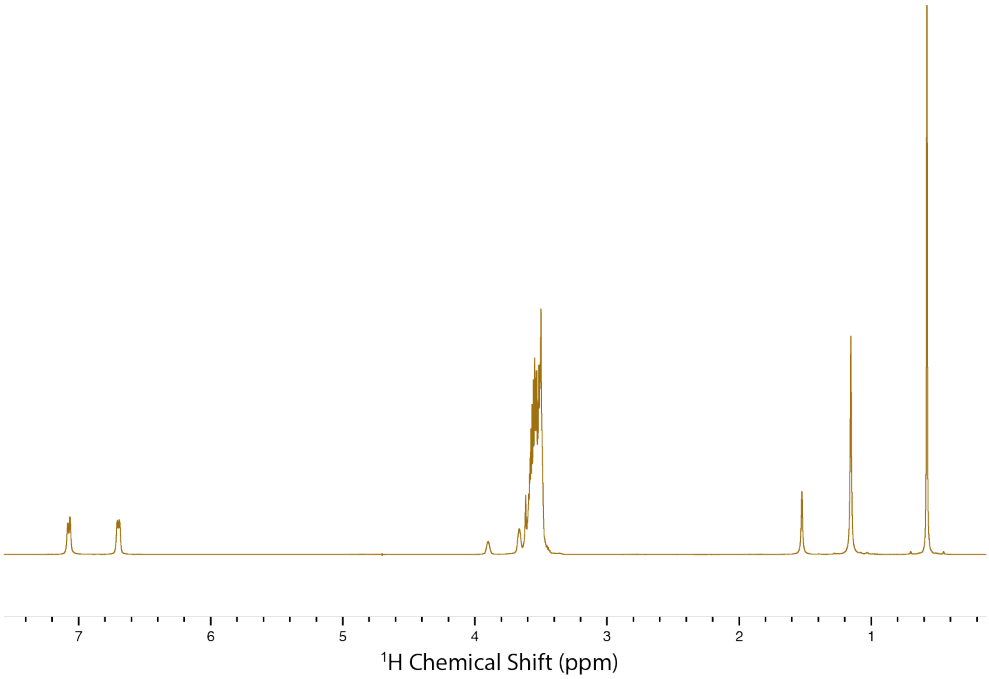


Figure 5.15: Ultracentrifugation of various liposomes.

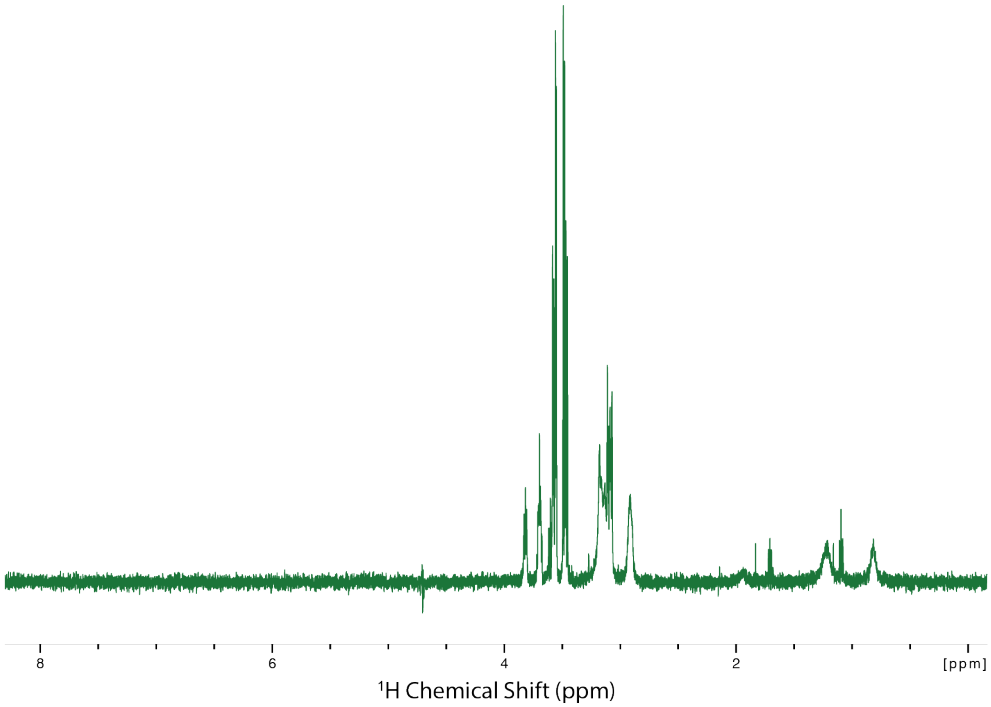
5.4.2 Reconstitution of PDGFβR-TM-20 into lipid vesicles

After PDGFβR-TM-20 was purified and cleaved as per the aforementioned protocol, the detergent was removed and the protein reconstituted into liposomes using Bio-Beads^{TM}SM-2 Resin as described in section 2.10.2. The equilibrium and kinetic aspects of reconstitution through the use of Bio-Beads^{TM}has been extensively studied (Holloway, 1973; Lévy et al., 1990) and its efficiency in removing Triton X-100 has been established (Allen et al., 1980). Furthermore, removal of detergent using Bio-Beads also allowed us to control the size of the liposomes formed, since their size is strongly dependent on the rate of detergent removal (Lévy et al., 1990).

To ensure that all detergent was removed following treatment with Bio-Beads, ^{1}H NMR spectra were acquired by solution-state NMR before and after detergent removal. Particular attention was paid to characteristic peaks of Triton X-100 such as the two aromatic peaks near 7 ppm (Figure 5.16a). Peaks arising from the buffer were also considered by acquiring ^{1}H NMR spectra of the buffer on its own and in the presence of detergent-solubilised lipids. As the detergent is removed and liposomes are formed, the turbidity of the solution increases and was monitored to check the rate of liposome formation. The ^{1}H NMR spectrum of the final suspension after the removal of the beads revealed the absence of Triton X-100 detergent peaks (Figure 5.16b).



(a) ^{1}H NMR spectrum of Triton X-100 micelles in water



(b) ^{1}H NMR spectrum after the removal of Triton X-100

Figure 5.16: Removal of Triton X-100 monitored using ^{1}H NMR spectra on a 500 MHz spectrometer at room temperature. Panel (a) shows the spectrum before removal and (b) shows the spectrum after removal.

5.4.3 Circular Dichroism (CD) spectroscopy

The secondary structure of PDGFβR-TM-20 was measured using CD experiments. CD is a commonly used biophysical technique for the determination of secondary structure of proteins and peptides and detects the differential absorption of circularly polarised light. Certain spectral signatures readily indicate an α-helical fold, and fitting software can be used to determine the relative percentages of α-helix, β-sheet and random-coil in the sample.

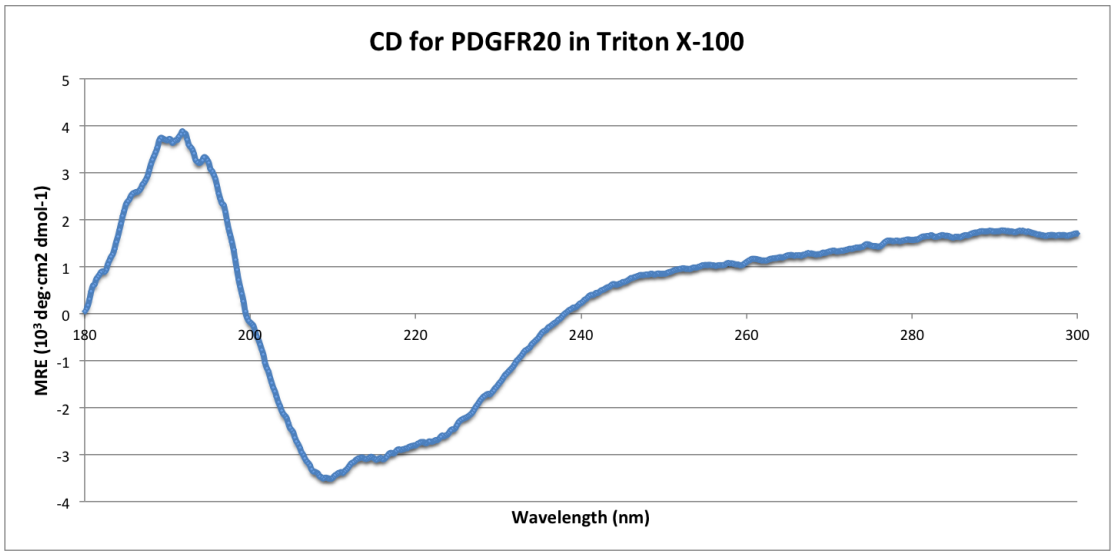
Purification was performed in a 20 mM HEPES buffer, but this buffer led to scattering at wavelengths below 200 nm. A literature search revealed that sodium phosphate, Tris and sodium acetate do not produce scattering (Kelly et al., 2005). In light of this, the protein samples were exchanged to 20 mM sodium phosphate buffer, pH 7.5 using a PD10 column. CD measurements were then carried out as detailed in section 2.11.

Initial CD characterisation of PDGFβR-TM-20 was performed in Triton X-100 and can be seen in Figure 5.17a. Visual inspection reveals the presence of bands at 210 nm and 222 nm indicative of predominant α-helical structure. Bands at these wavelengths are also observed when PDGFβR-TM-20 is embedded in a POPC bilayer, as can be seen in Figure 5.17b.

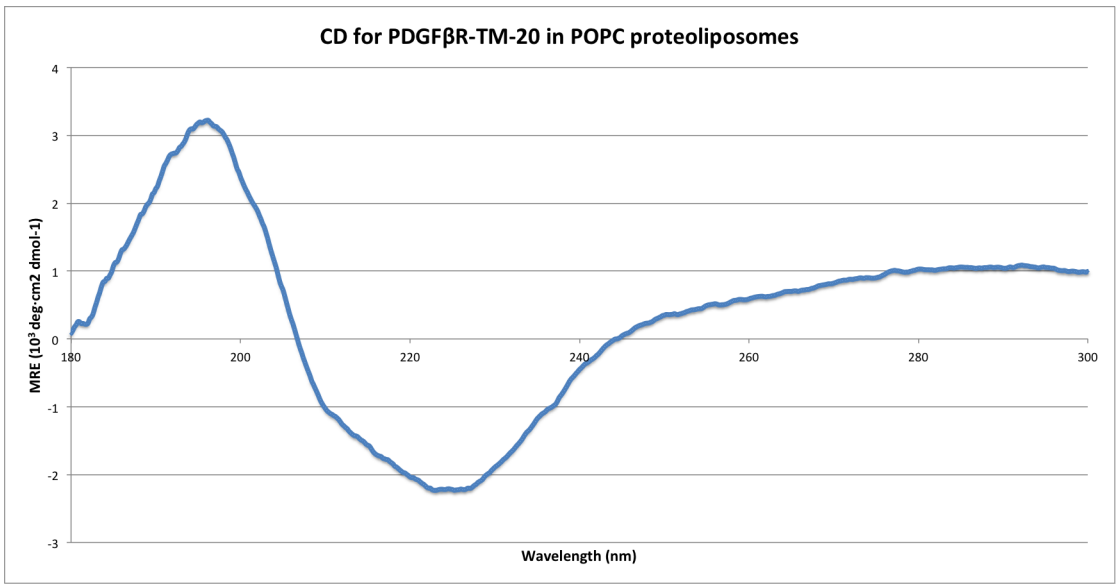
The CD data were also analysed using a fitting software (DichroWeb, CDSSTR) (Whitmore and Wallace, 2004) which confirms a predominantly α-helical fold for PDGFβR-TM-20 in both Triton X-100 detergent and POPC bilayers. Table 5.1 shows the percentage distribution of secondary structure features obtained from DichroWeb using the K2D algorithm with reference set 4 (190-240 nm), where 74.9% of the protein is predicted to be in an α-helical conformation in Triton X-100 and 68.4% of the protein is predicted to be in an α-helical conformation in POPC bilayers. These results indicate that the protein is well-folded within POPC bilayers and suitable for further analysis through solid-state NMR.

Membrane mimetic	α-helix (%)	β-sheet (%)	Turns (%)	Unordered (%)
Triton X-100	74.9	0	9.4	14.6
POPC	68.4	0	13.9	17.7

Table 5.1: Secondary structure percentage distribution for PDGFβR-TM-20 in membrane mimetics according to DichroWeb.



(a) CD spectrum for PDGFβR-TM-20 in Triton X-100



(b) CD spectrum for PDGFβR-TM-20 in POPC liposomes

Figure 5.17: CD for spectra for PDGFβR-TM-20 in membrane mimetics. For both measurements, ∼0.25 mg/mL of the peptide in a buffer containing 20 mM sodium phosphate, 20 mM NaCl at pH 7.4) was used.

5.4.4 Labelling strategy and initial NMR results

A recently developed labelling strategy that relies only on the use of [1-^{13}C]glucose and [2-^{13}C]glucose (Loquet et al., 2010) has been used to successfully acquire a high-resolution NMR structure of type III secretion system needles in bacteria which had previously not been amenable to structural studies at the atomic level (Loquet et al., 2012; Demers et al., 2014). This sparse-labelling scheme greatly improves spectral resolution and many of the cross-peaks can be assigned to long-range distance restraints.

With this strategy in mind, a sample of PDGFβR-TM-20 was expressed with [1-^{13}C]glucose as the sole carbon source during expression. The protein was reconstituted into POPC liposomes (see section 2.10.2 for details) and then pelleted into an MAS ssNMR rotor for analysis using ^{13}C–^{13}C correlation experiments. The experimental parameters and set up for solid-state NMR experiments were optimised as for the Neu* peptide (section 3.7). After collecting an alanine CP spectrum for controls and optimising CP on [1-^{13}C] PDGFβR-TM-20 in POPC liposomes, 2D MAS ssNMR DARR experiments were performed.

A 2D ^{13}C–^{13}C DARR spectrum obtained using a 150 ms mixing time is shown in Figure 5.18. The strong signals on the diagonal are due to either the labelled amino acids in PDGFβR-TM-20 or from natural abundance ^{13}C present in POPC lipids. A limited number of cross-peaks are also observed on either side of the diagonal, in both the aliphatic and carbonyl regions of the spectrum. However, these peaks were significantly fewer in number than expected. Furthermore, the signal to noise ratio was so low that the experiment had to be performed for 6 days in order for signals to be observable above the noise. This poor S/N could be due to inefficient reconstitution of the protein into membrane bilayers or dynamics and conformational exchange within the bilayer on an intermediate timescale.

Unfortunately, it was not possible to obtain any distance restraints using this spectrum, and the low signal to noise ratio made further experiments infeasible. Troubleshooting needs to be performed before proceeding further; however, this could

not be done due to the limited timeframe of this project.

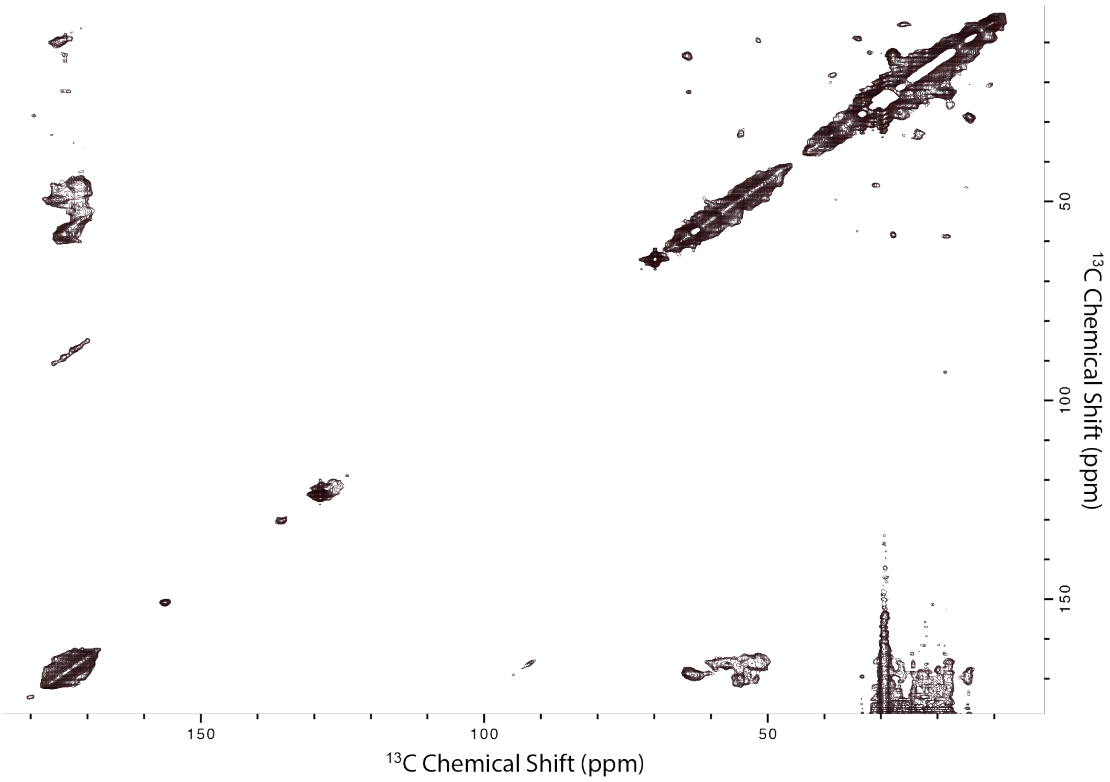


Figure 5.18: 2D ^{13}C–^{13}C DARR correlation spectrum of PDGFβR-TM-20 in POPC liposomes, acquired over 6 days with a mixing time of 150 ms. Spectra were recorded at 500 MHz spectrometer with 8.5 kHz MAS at 10 °C with 552 scans.

5.5 Sample Preparation for Solution-State NMR

While solid-state NMR is an invaluable technique in terms of probing the structure and dynamics of membrane proteins in bilayers, more traditional solution-state NMR methods can still yield crucial structural and mechanistic insights into membrane protein function and interactions, bypassing the difficulties that stymie ssNMR analyses. As was explained in section 4.1, PDGFβR interacts with E5 from bovine papillomavirus but the molecular details of this interaction are poorly understood. Studying this interaction in detergent micelles would yield the first direct evidence that E5 binds to the TM of PDGFβR and shed new light on the binding interface. In the first instance, it was decided to continue investigations with the His-tag attached to the protein construct as that would significantly reduce sample preparation time and ease further investigations.

5.5.1 Purification of PDGFβR-TM-20

In pursuit of information relating to E5/PDGFβR-TM interaction, a protocol for the purification of PDGFβR-TM-20 was developed to prepare samples in detergent micelles (or mixed micelles) for solution-state NMR analysis. This has been explained in detail in section 2.6.3 and involved purifying the protein in a chaotropic agent such as urea, followed by dialysis and an additional HPLC purification step.

SDS-PAGE gels suggested that the protein was largely pure post-dialysis (Figure 5.19), however, initial NMR screening revealed poor solution behaviour leading to broad NMR peaks in the ^{1}H–^{15}N HSQC spectrum (Figure 5.20a). An additional HPLC step lead to reduction of noise and improved NMR peak shape and intensity (Figure 5.20b). Furthermore, it was found that the protein powder obtained after HPLC dissolved more readily circumventing dissolution issues encountered when dissolving the protein powder directly after dialysis.

HPLC purification was performed as detailed in section 2.6.3 with a semi-preparative C4 reverse-phase HPLC column. The chromatographic profile shown in Figure 5.21

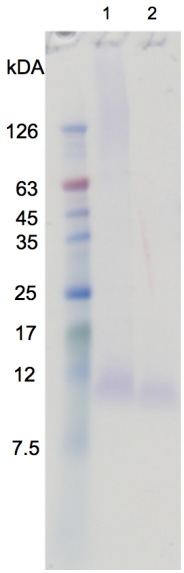
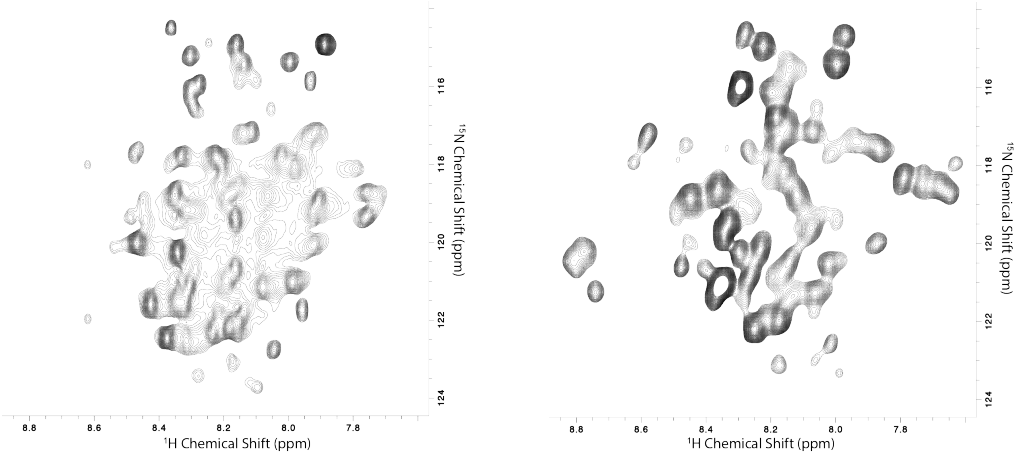


Figure 5.19: SDS-PAGE gel of PDGFβR-TM-20 before and after HPLC purification. Lane 1: before HPLC purification, Lane 2: after HPLC purification.



(a) HSQC before HPLC **(b)** HSQC after HPLC

Figure 5.20: ^{1}H–^{15}N HSQC spectra before and after the introduction of an additional HPLC step in the purification process. Data was acquired at 37 ∘C on a 750 MHz spectrometer. PDGFβR-TM-20 was dissolved in a 20 mM sodium phosphate buffer with 100 mM DPC detergent at pH 6.8.

was obtained, measuring absorbance at 280 nm.

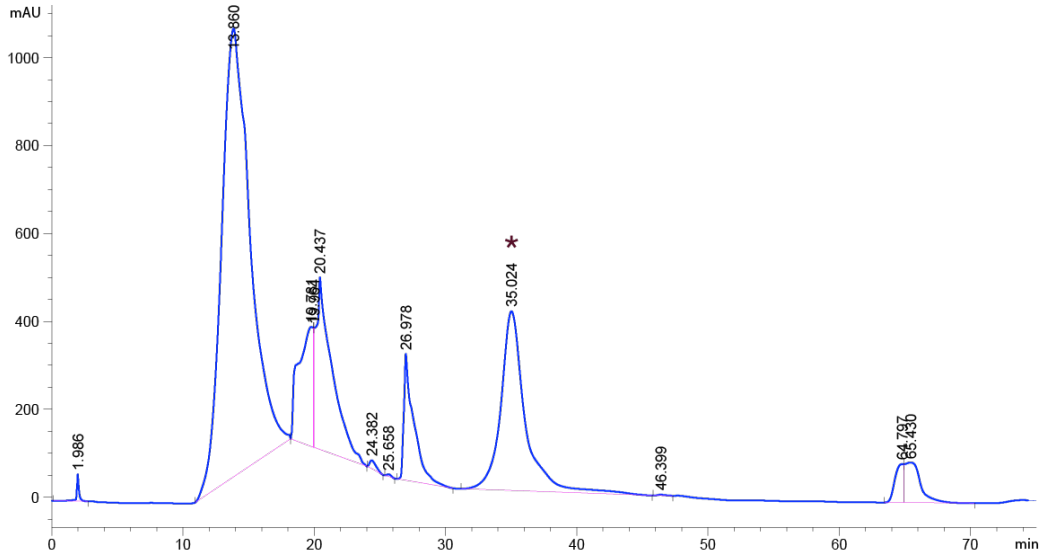
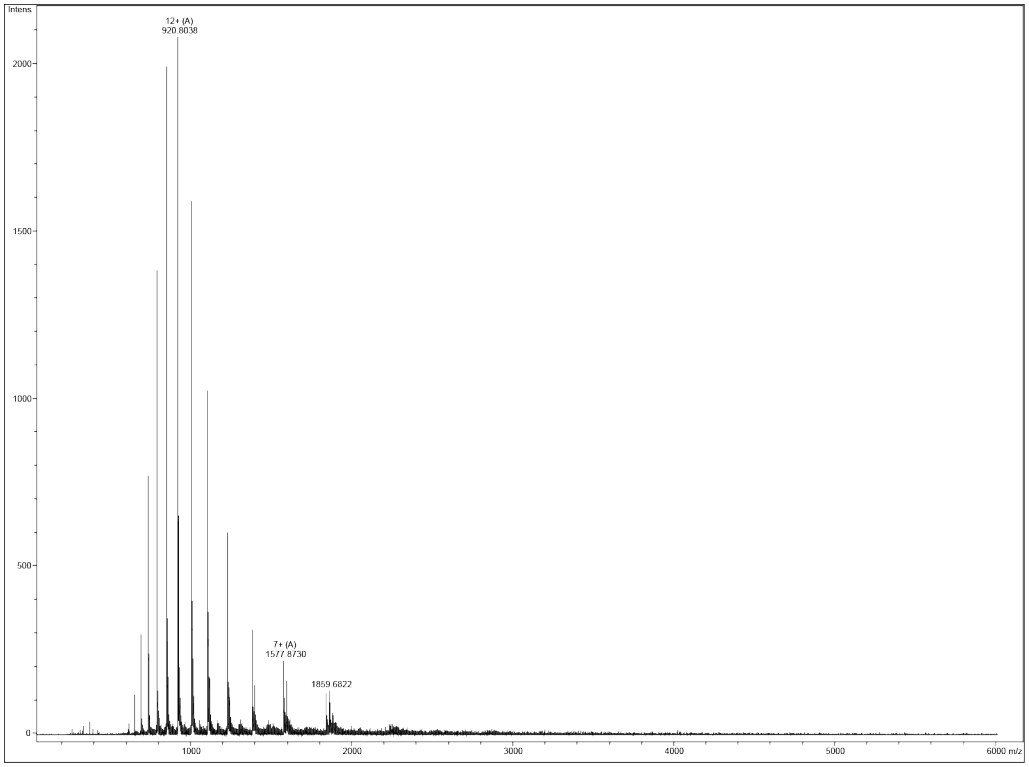


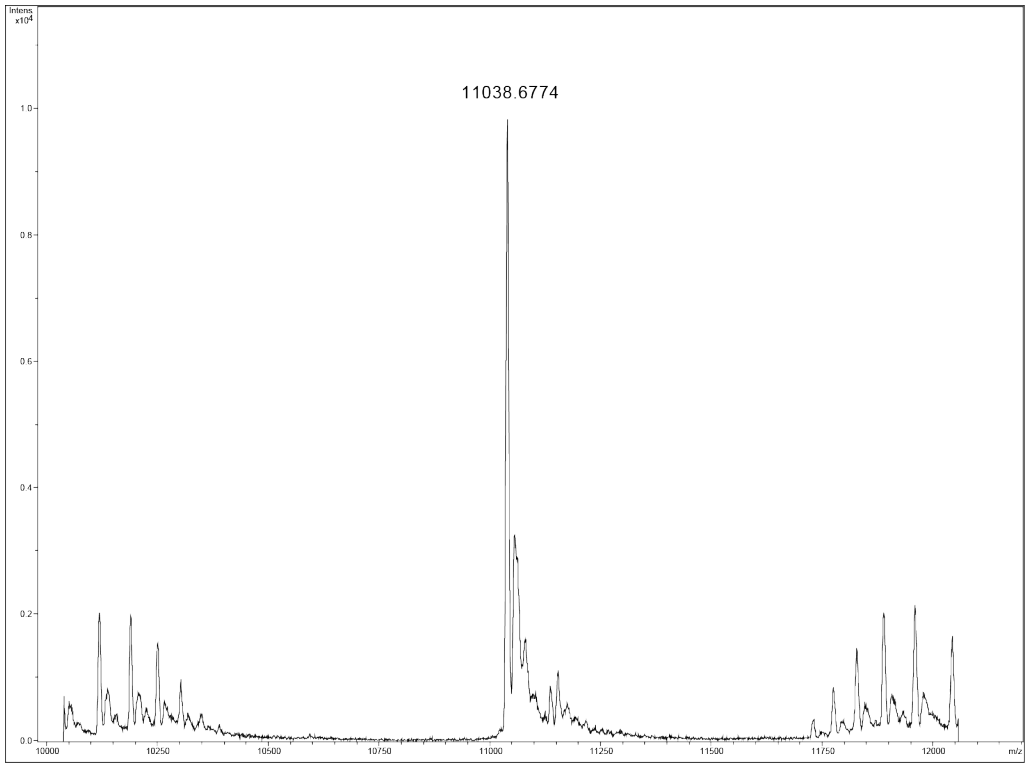
Figure 5.21: HPLC chromatogram for the purification of PDGFβR-TM-20. Peak containing the protein is marked with an asterisk (*).

Mass spectrometry was performed on HPLC fractions obtained corresponding to major peaks in the chromatogram. The peak at 35 minutes (marked with (*)) was found to contain PDGFβR-TM-20. Figure 5.22a shows the mass spectrum of ^{15}N labelled PDGFβR-TM-20 with various charge states and Figure 5.22b shows the corresponding deconvoluted spectrum with a peak for PDGFβR-TM-20 at a mass of 11038.68 Dam which agrees well with the expected mass of 11037.78 Da.

All subsequent NMR samples were prepared by directly dissolving PDGFβR-TM-20 into the detergent buffer solution up to the required concentration. The concentrations were measured using Beer-Lambert law.



(a) Mass spectrum of PDGFβR-TM-20



(b) Deconvoluted mass spectrum of PDGFβR-TM-20

Figure 5.22: Assessment of purified PDGFβR-TM-20 through ESI-MS. Pooled fractions from the HPLC purification were checked for the existence of the protein and it's purity. The top panel (a) shows the mass spectrum with multiple charge states of the protein. The bottom panel (b) shows the deconvoluted spectrum with a major peak corresponding to the theoretical mass of PDGFβR-TM-20 protein.

5.5.2 Purification of E5

A peptide corresponding to the TM domain of E5:

KKKFLGLVAAMQLLLLLFLLLFFLVYWDHK

was ordered from a commercial manufacturer, and provided in a crude form containing truncated products and protecting groups from the solid-phase synthesis. It was purified using reverse-phase HPLC (RP- HPLC) as detailed in section 2.6.3 and then analysed by mass spectrometry.

Elution fractions corresponding to major peaks in the HPLC chromatogram (Figure 5.23) were collected and pooled. The peak at around 46.3 minutes (marked with (*)) was found to contain E5. This was deduced from the mass spectrometry data (Figures 5.24a & 5.24b) which revealed a peak at 3662.2 Da in the deconvoluted spectrum, which is in good agreement with the theoretical mass of 3664.60 Da.

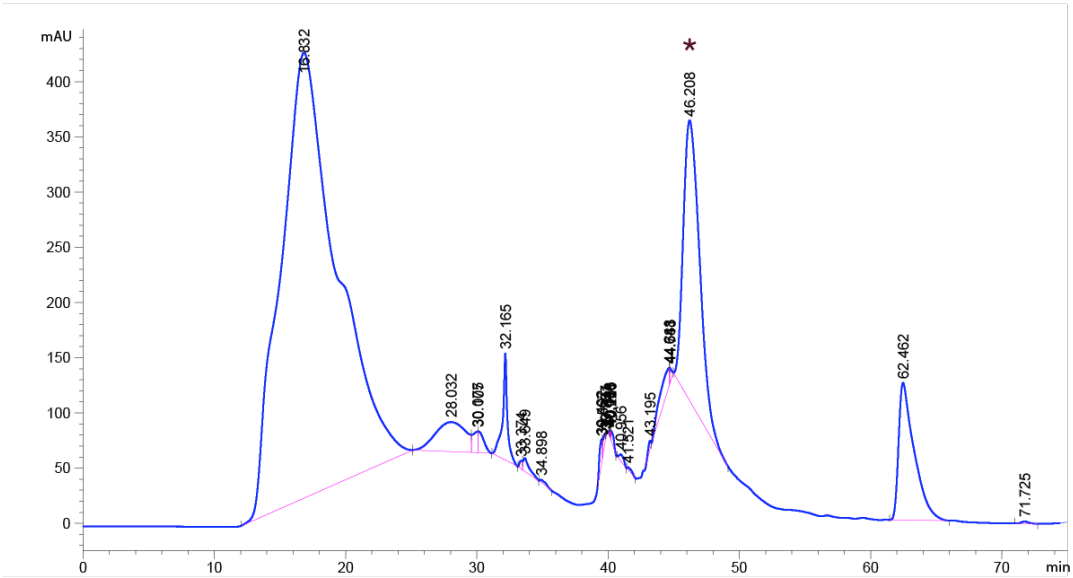
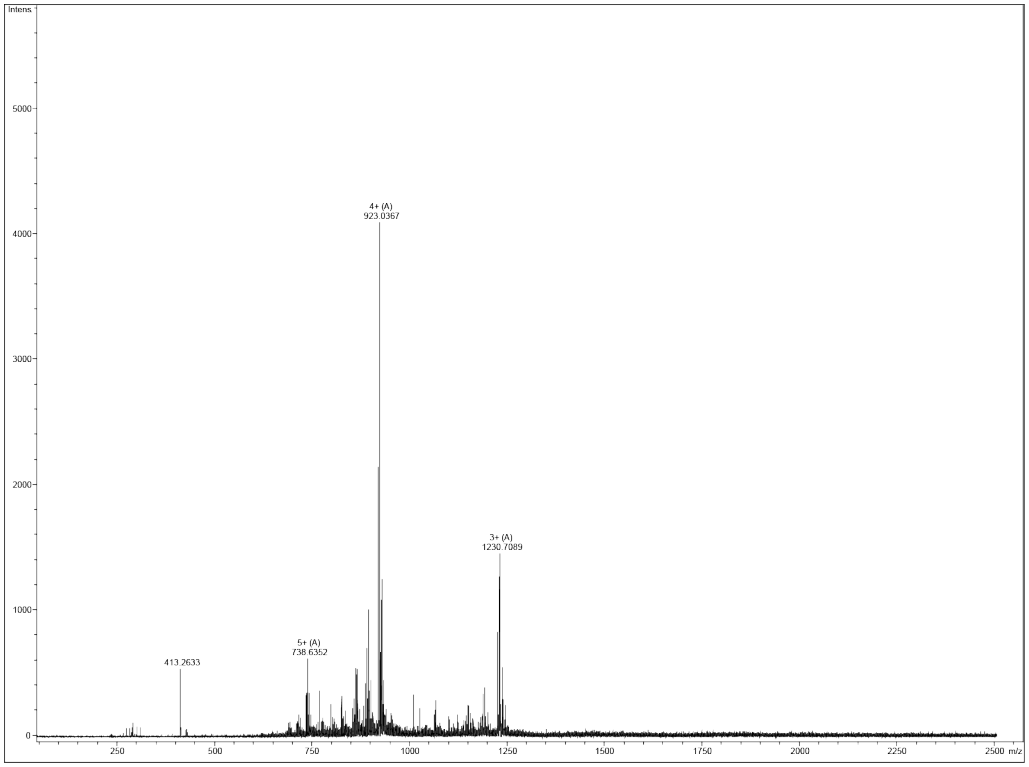
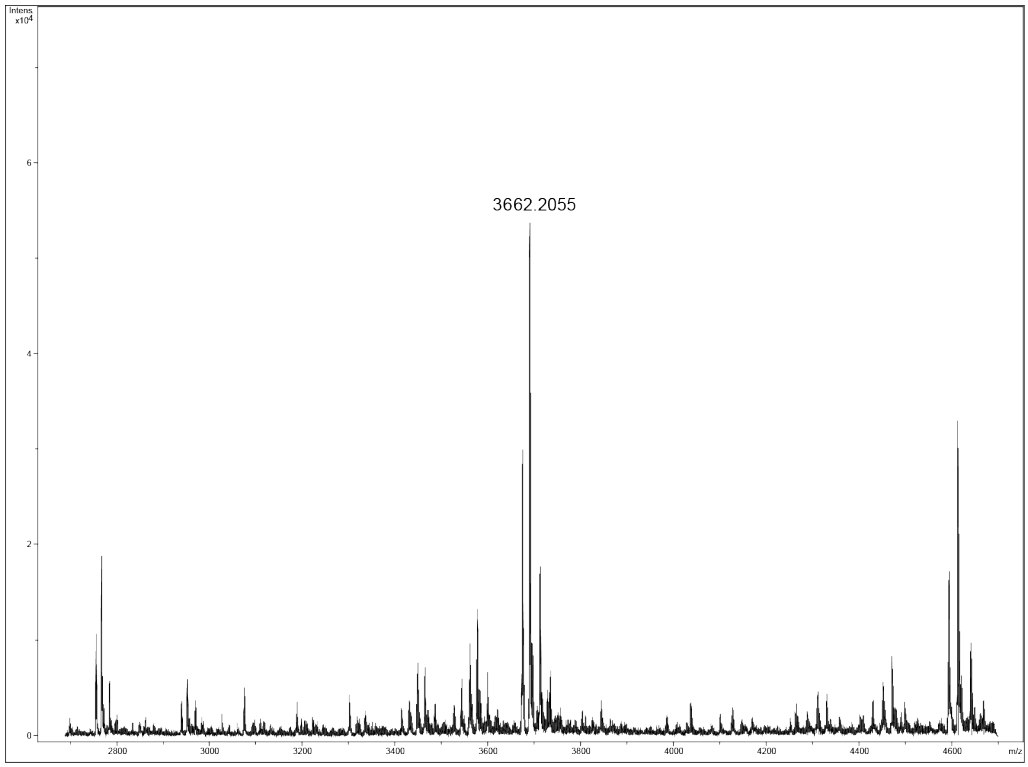


Figure 5.23: HPLC chromatogram for the purification of E5

The pure E5 peptide was stored in TFE. The concentration was measured using absorbance at 280 nm and Beer-Lambert Law. E5 was then titrated into the PDGFβR sample by first drying it under a stream on nitrogen such that a thin film was formed. This film was then dissolved in a solution containing PDGFβR-TM-20 in DPC/DPPC mixed micelles.



(a) Mass spectrum of E5



(b) Deconvoluted mass spectrum of E5

Figure 5.24: Assessment of purified E5 through ESI-MS. Pooled fractions from the HPLC purification were checked for the existence of the peptide and it's purity. The top panel (a) shows the mass spectrum with multiple charge states of the protein. The bottom panel (b) shows a major peak corresponding to E5.

5.6 Summary and Conclusion

This chapter began by detailing the modifications that were carried out to purify PDGFβR-TM-20 which was expressed using an existing plasmid construct with an (rEK) enterokinase cleavage site. This was important because even though a protocol existed in the Dixon lab for the purification of unlabelled PDGFβR-TM-20, it needed to be ensured that yields were maximised when the protein was expressed in expensive labelled media. The changes involved switching the chaotropic agent to urea, changing the volumes of elution and wash buffers and the concentration of imidazole in them. Also, the choice of detergent was changed to Triton X-100 as it could easily be removed using Bio-Beads for reconstitution into lipid bilayers.

It was subsequently observed that despite optimising the cleavage conditions, it would not be economically feasible to use rEK for cleavage for the quantities of protein required for NMR. Additionally, the construct contained an uncleavable C-terminal His-tag whose removal would require genetic manipulation. In light of these difficulties, a new construct was created using TOPO cloning with a His-tag cleavable using TEV protease. The construct was created exactly as per requirements and TEV protease was chosen because protocols existed in the lab to produce it in large quantities in an economically efficient manner.

However, it soon became apparent that the self-cleaving MBP which is produced when expressing TEV protease cannot be removed from the sample using any of the purification methods that were tried (section 5.3.4). A GFP-tagged TEV protease was then acquired which resolved this issue and samples could finally be made for NMR analyses.

A protocol for the reconstitution of PDGFβR-TM-20 into liposomes using Bio-Beads was adopted which monitored the removal of detergent using solution-state NMR. Insertion and secondary structure classification was studied using CD which revealed a highly helical structure of the protein within Triton X-100 detergent and POPC liposomes.

Two-dimensional ^{13}C–^{13}C DARR correlation spectra were acquired of the liposomes after inserting PDGFβR-TM-20 labelled with [1-^{13}C]glucose. Peaks were observed in the aliphatic as well as carbonyl regions of the spectrum. However, not only was the signal to noise ratio very low, but the number of peaks were significantly lower than expected such that meaningful distance restraints could not be obtained. This necessitated a re-examination of several aspects of the sample preparation process. It may be that instead of spontaneously inserted into the membrane upon detergent removal, a certain percentage of the protein is aggregating. Approaches to circumvent this could involve changing the bilayer composition, or the length of the juxtamembrane region. A more traditional uniform labelling scheme could also be adopted. Due to time-constraints, this problem could not be immediately addressed.

The last section of this chapter dealt with the sample preparation methods for the investigation of E5/PDGFβR-TM interactions using solution-state NMR. A double purification method involving IMAC and HPLC were adopted to obtain pure PDGFβR-TM-20 for analyses. The following chapter will address in detail the solution-state NMR experimental optimisations performed and the results obtained from these studies.

Chapter 6

A Solution-State NMR study of E5/PDGFβR Transmembrane Interactions

6.1 Introduction

Solid-state NMR is a very powerful technique that can be employed to investigate structural properties and interactions in a lipid bilayer, however, current data acquisition methods render it far less sensitive than its solution counterpart. Where cutting-edge methods are being developed for structural assignments of proteins in ssNMR, assignment methods using 3D NMR are well-established in the solution state. In light of this, it was decided to investigate E5/PDGFβR interactions using solution-state NMR as a complementary method.

Solution-state NMR requires fast tumbling on the NMR timescale in order to average anisotropic NMR interactions (e.g. CSA), therefore large biomolecules or complexes $>$ 50 kDa pose significant difficulty in yielding adequate resolution and signal to noise ratio. As such, liposomes do not constitute a suitable membrane-mimetic for studying membrane embedded proteins using solution-state NMR. Detergents, on the other hand, have been used to study membrane proteins *in vitro* for the past 40 years (Helenius and Simons, 1975; Tanford and Reynolds, 1976) and most NMR structures of membrane proteins have been solved in detergent solutions (Kang and Li, 2011; Warschawski, 2013).

This chapter will introduce NMR experiments performed on uniformly ^{15}N labelled PDGFβR-TM-20 in detergent micelles and establish a suitable system for acquiring more complex 3D experiments. A set of resonances was hypothesised to belong to the TM region of PDGFβR-TM-20 through Hydrogen–deuterium exchange experiments. PDGFβR-TM-20/E5 interactions were explored by titrating E5 into the PDGFβR-TM-20 mixed micelle solution and the implications of the results obtained will be described.

6.2 The Heteronuclear Single Quantum Coherence (HSQC) Experiment

First described in 1980 (Bodenhausen and Ruben, 1980), the HSQC experiment has found its place as one the most important, and most frequently used, experiments in solution-state biomolecular NMR. It is usually the first heteronuclear experiment to be performed on proteins, and the spectrum acquired forms a 'fingerprint' because there is a peak for each NH residue in the protein backbone (apart from prolines). The HSQC spectrum allows us to assess whether higher dimensional experiments are feasible, and whether experiments exploiting the use of expensive carbon labels are economically viable. In the case of large proteins or complexes, the need for deuteration can also be evaluated.

The HSQC experiment is initiated by transferring magnetisation from the proton to an attached ^{15}N (or ^{13}C) nucleus via INEPT (Insensitive Nuclei Enhanced by Polarisation Transfer). The magnetisation is subsequently transferred back to the proton after a certain delay (termed t_1) for detection. A 2D HSQC spectrum is produced by collecting a series of experiments with an increasing t_1 delay (see section 1.5.6 for details).

This experiment was heavily used in this work to optimise PDGFβR-TM-20 sample quality. Projections of other 3D experiments are also compared with the HSQC spectrum to ensure no sample degradation had occurred. Additionally, the effect of

deuterium exchange and chemical shift perturbations upon the addition of E5 ligand were monitored using the HSQC experiment.

All experiments were conducted on a 700 MHz spectrometer as detailed in section 2.14.

6.3 Optimisation of HSQC experiments

6.3.1 Analysis of [*U*−^{15}N]-PDGFβR-TM-20 in trifluoroethanol (TFE)

Prior to performing detailed analyses in detergent solutions, a ^{1}H–^{15}N HSQC spectrum was acquired for PDGFβR-TM-20 solubilised in deuterated TFE (TFE-d_3) (Figure 6.1). Since TFE promotes secondary structure formation (Roccatano et al., 2002) while destabilising intermolecular helix–helix interactions and hence disrupting oligomerisation (Luo and Baldwin, 1997), the homogenous population, high signal to noise and narrow linewidths allowed us to get an idea of dispersion and confirm incorporation of ^{15}N labels.

A total of 112 peaks (including side-chain peaks) were expected in the PDGFβR-TM-20 HSQC spectrum, and a total of 110 were observed (Figure 6.1). Based on their unique position in the spectrum, we identified all 6 Gly residues (expected from sequence) in the PDGFβR-TM-20 spectrum (highlighted in a blue box in Figure 6.1). Similar observations were made regarding the side-chain peaks of Trp, Arg, Asn and Gln and are summarised in Table 6.1.

Category	Peaks expected	Peaks observed
Gly backbone	6	6
Gln and Asn side-chain	6	6
Trp side-chain	3	3
Arg side-chain	2	2

Table 6.1: List of identifiable peaks in an HSQC spectrum of PDGFβR-TM-20 in TFE-d_3 by amino acid type.

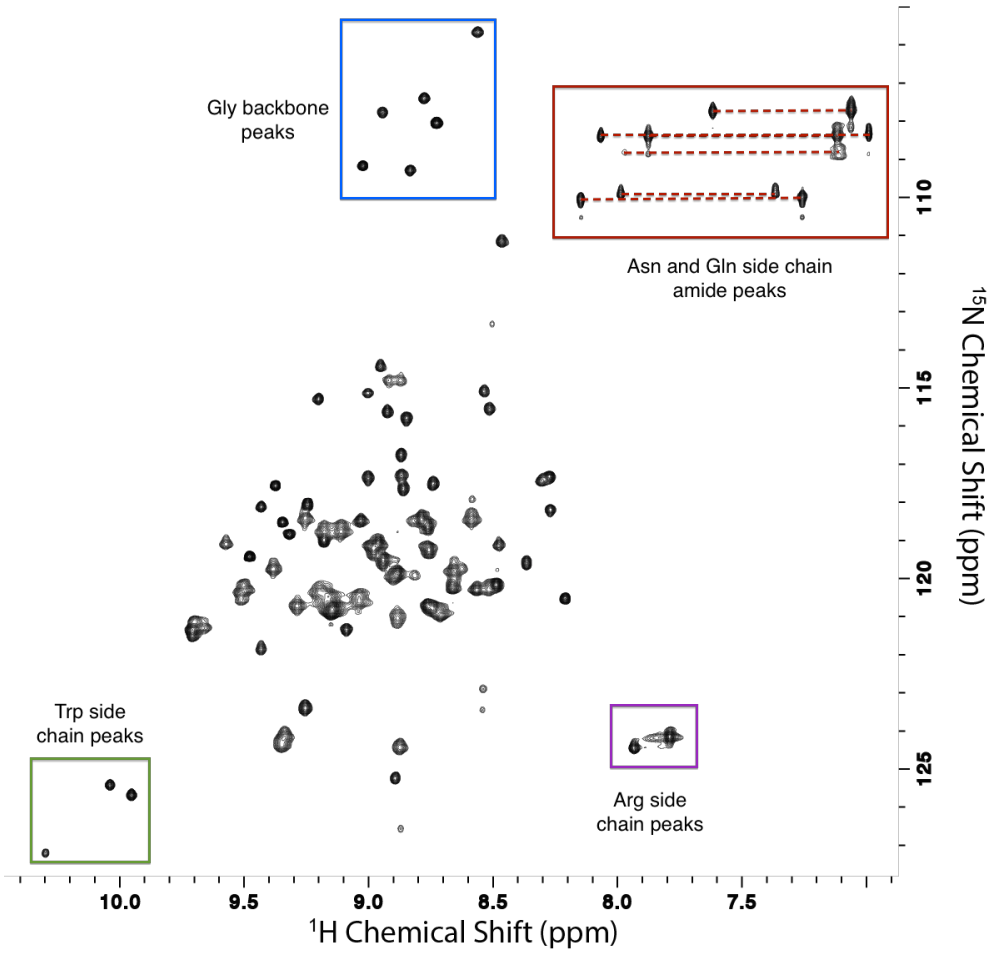


Figure 6.1: A ^{1}H–^{15}N HSQC spectrum of U–^{15}N-PDGFβR-TM-20 dissolved in TFE-d$_3$ on a 700 MHz spectrometer. Blue box identifies peaks which are present in a region characteristic of Gly peaks. Similarly, red box anticipates Asn and Gln side-chain peaks, green box anticipates Trp peaks and purple box anticipates Arg peaks.

6.3.2 Detergent screen

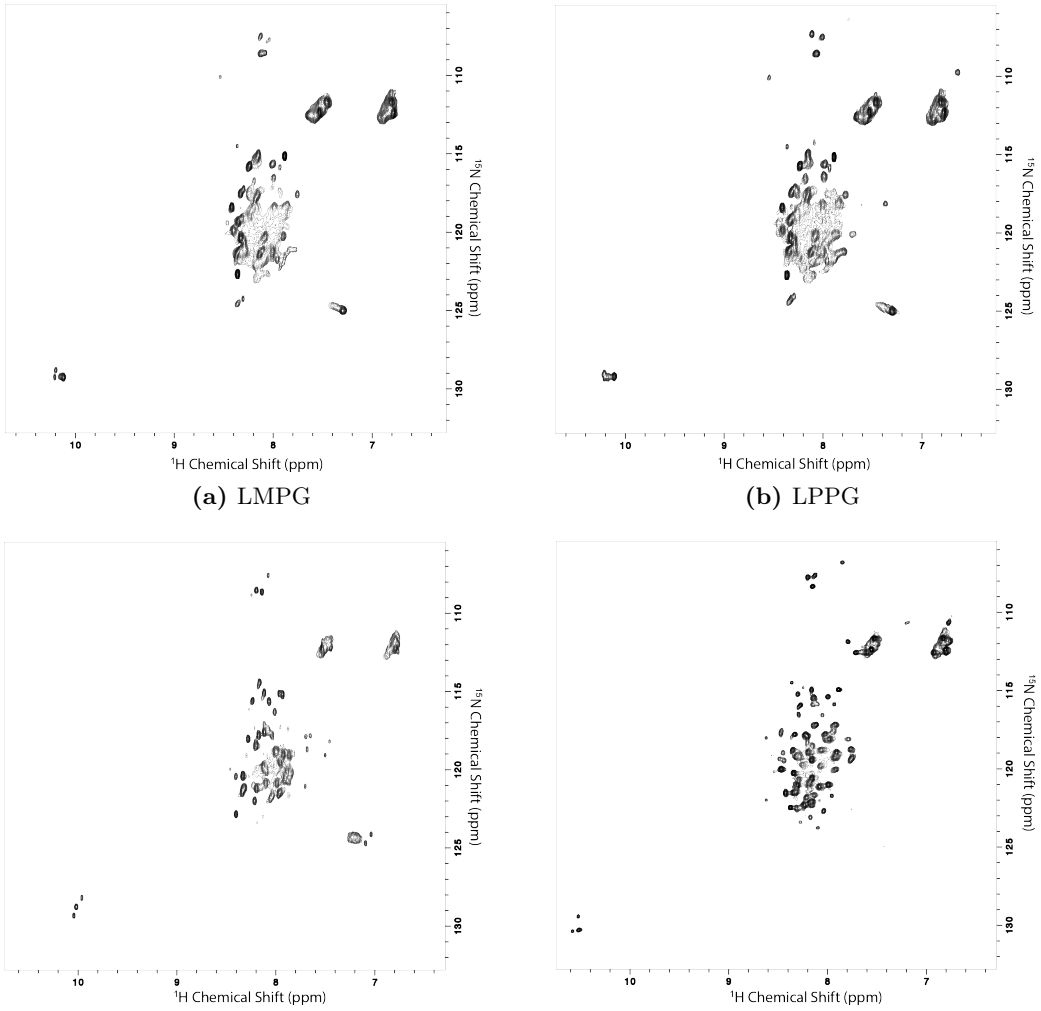
Although detergent micelles are the most commonly used solubilising agents for solution NMR studies, there are no generally applicable rules which dictate the choice of detergent (Poget and Girvin, 2007). Nonetheless, the choice of detergent is crucial as a determining factor of spectral quality (Krueger-Koplin et al., 2004). A total of 4 detergents were tested to determine which would yield the expected 112 peaks in the HSQC spectrum (Figure 6.2). It was immediately clear that the detergent micelles which offered the best spectral resolution were those formed by the zwitterionic detergent n-dodecylphosphocholine (DPC). The lysophos-detergents 1-myristoyl-2-

hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) and 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LPPG) along with the anionic detergent sodium dodecyl sulfate (SDS) did not yield adequate signal to noise ratio. Subsequent NMR spectra were hence collected by solubilising the protein in DPC, as the spectrum acquired even without further optimisation was deemed of sufficient quality (by checking the number of peaks and linewidths) for monitoring of chemical shift perturbations upon the addition of E5 ligand, which would provide useful information on binding. It is estimated the size of the protein/micelle complex is about 30 kDa.

6.3.3 Temperature

Slow molecular reorientation or tumbling can lead to increased linewidths and spectral overlap via fast relaxation of transverse magnetisation (short T_2). A simple strategy to minimise this effect in the large membrane protein/micelle complexes is acquire data at elevated temperatures.

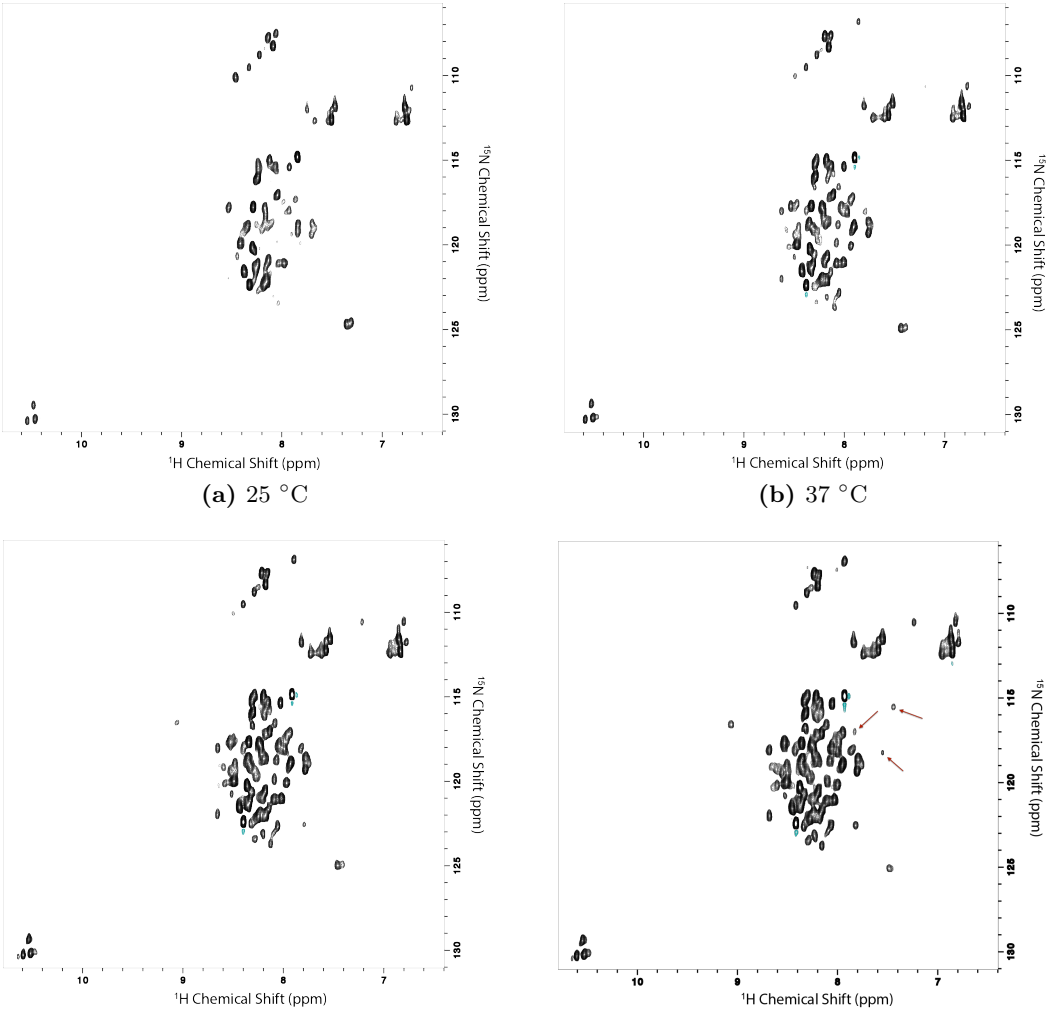
Typically, temperatures above 30 °C are preferred for NMR spectroscopy of membrane proteins (Fernández and Wüthrich, 2003). A temperature range of 25–45 °C was tested, with measurements collected at 25 °C and then at 4 °C increments between 37 °C and 45 °C. These spectra can be seen in Figure 6.3, and suggest that there is a dramatic improvement in spectral quality when increasing temperature from 25 °C to 37 °C. It continues to improve gradually up to 45 °C when several new peaks start to appear (highlighted with red arrows). The protein was stable at a temperature of 45 °C for at least a week, hence all subsequent experiments were performed at 45 °C.



(a) LMPG **(b)** LPPG

(c) SDS **(d)** DPC

Figure 6.2: ^{1}H–^{15}N HSQC spectra of [U–^{15}N]-PDGFβR-TM-20 solubilised in different detergent micelles. LMPG (a) and LPPG (b) were used at a concentration of 50 mM whereas SDS and DPC are used at a concentration of 100 mM. All experiments are conducted on 0.1 mM of protein at 310 K in a 50 mM sodium phosphate buffer with 20 mM NaCl and pH 6.0.



(a) 25 °C **(b)** 37 °C

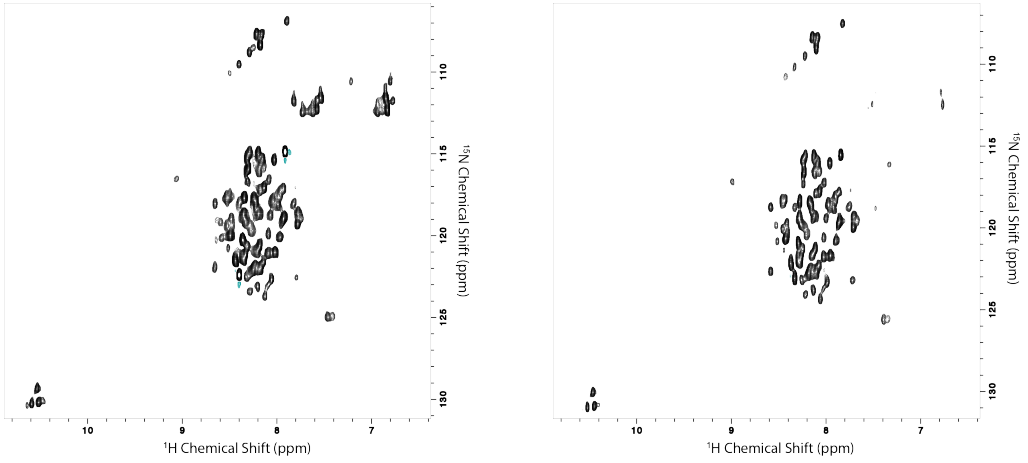
(c) 41 °C **(d)** 45 °C

Figure 6.3: ^{1}H–^{15}N HSQC spectra of [U–^{15}N]-PDGFβR-TM-20 at various temperatures. All experiments are conducted on 0.1 mM of protein in a 50 mM sodium phosphate buffer with 20 mM NaCl and pH 6.0.

6.3.4 The transverse relaxation optimised spectroscopy (TROSY) effect

TROSY-based NMR (Pervushin et al., 1997) methods have been utilised for a number of years to successfully overcome the 30–40 kDa protein size limit encountered when using traditional NMR pulse sequences. Application of TROSY-based methods has allowed acquisition of assignable spectra of proteins with a size of about 150 kDa (Salzmann et al., 2000; Tugarinov et al., 2005). It has also been observed that improvements in ^{1}H–^{15}N correlation spectra are also observed using the TROSY-effect, even in the absence of deuteration (Kim et al., 2009).

Experiments using traditional HSQC and TROSY pulse programmes were performed on the same [U–^{15}N]-PDGFβR-TM-20 sample in DPC detergent (Figure 6.7a & 6.4b). However, no significant improvements were observed in the TROSY spectrum in terms of the number of backbone peaks or linewidths.

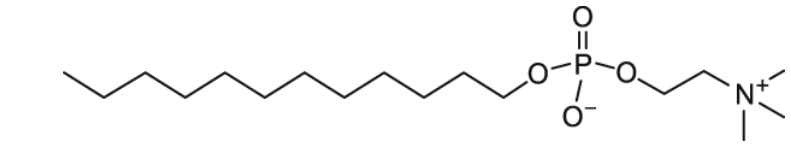


(a) 2D-[^{1}H,^{15}N]-HSQC **(b)** 2D-[^{1}H,^{15}N]-TROSY

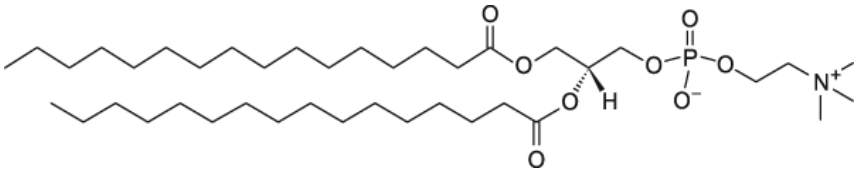
Figure 6.4: ^{1}H–^{15}N correlation spectra with and without the TROSY-effect.

6.3.5 Addition of DPPC

Recently, it was reported that the addition of 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) to detergents has provided a more native-like lipid environment around the protein without significant loss in signal to noise (Ortega-Roldan et al., 2015; Brady et al., 2015). In light of these observations, it was decided to add DPPC to DPC micelles at a q of 0.1, where $q = \frac{[\text{lipid}]}{[\text{detergent}]}$. The chemical structures of DPC and DPPC are shown in Figures 6.5a & 6.5b respectively.



(a) DPC



(b) DPPC

Figure 6.5: Chemical structures of (a) DPC and (b) DPPC.

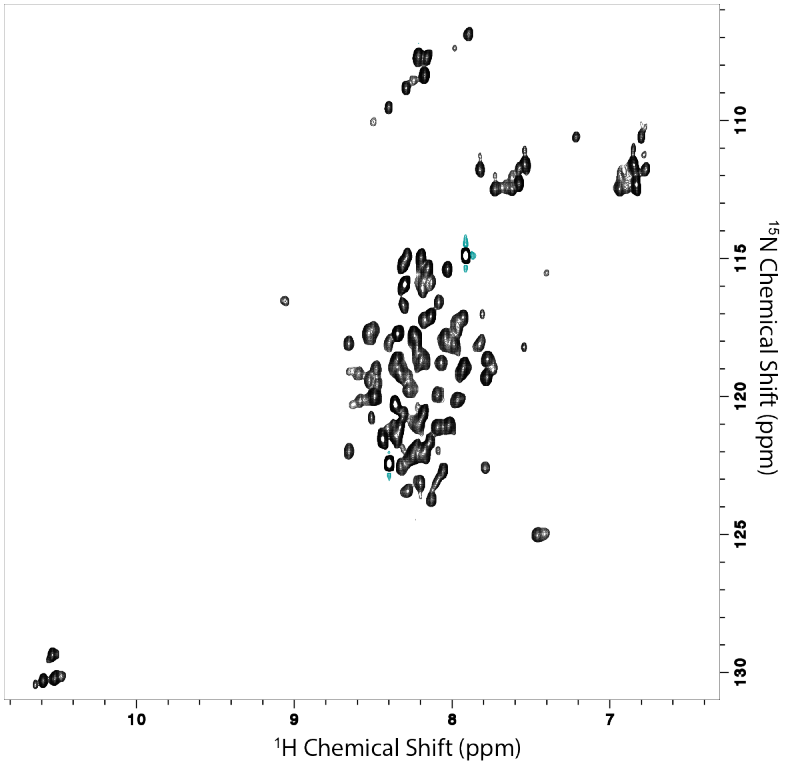
It can be seen through a comparison of Figures 6.6a & 6.6b that a large number of resonances remain unperturbed upon addition of DPPC. This is not surprising because the lipids would be expected to preferentially solvate the transmembrane helix and any chemical shift changes that are seen would be expected around the lipid–buffer–protein interface where the phosphocholine headgroup of DPC is replaced by the phosphatidylcholine headgroup of DPPC. Only minor changes (if any) are expected in the central regions of the transmembrane domain because the acyl chains of the lipid and the detergent are chemically similar. Similarly, since the chemical environment around the soluble juxtamembrane regions remains unaltered, few changes in chemical shifts would be expected.

Adding DPPC to the detergent solution yielded three important benefits. First, it was observed that peak widths throughout the spectrum were reduced leading to greater resolution. This is likely due to the lipids granting the protein a greater con-

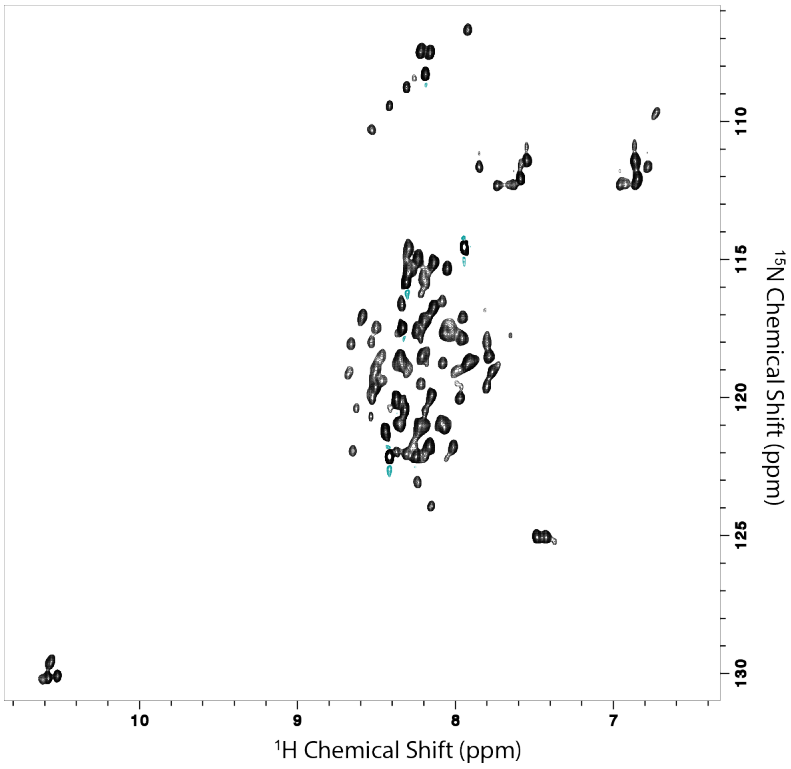
formational stability. This is particularly evident in the case of Trp side-chain peaks. When solubilised in DPC, two of the three Trp side-chain peaks exhibited 'splitting' where two peaks were observed for a single amide due to slow chemical exchange between two conformational states on the NMR timescale (Figure 6.7a).

Additionally, more peaks that can be attributed to residues in the TM domain were observed after the addition of DPPC. This was deduced after conducting D_2O exchange experiments with and without the addition of DPPC as explained in section 6.4.

Finally, in DPC, the maximum concentration of PDGFβR-TM-20 that could be solubilised before the chemical shifts started to change was around 0.3 mM. Upon the addition of DPPC, the protein had a far greater conformational stability, and no major changes in chemical shifts were observed. This was particularly important as in order to do backbone assignments 3-dimensional NMR experiments need to be performed which require very high signal to noise ratios which can be aided using a high concentration of protein.

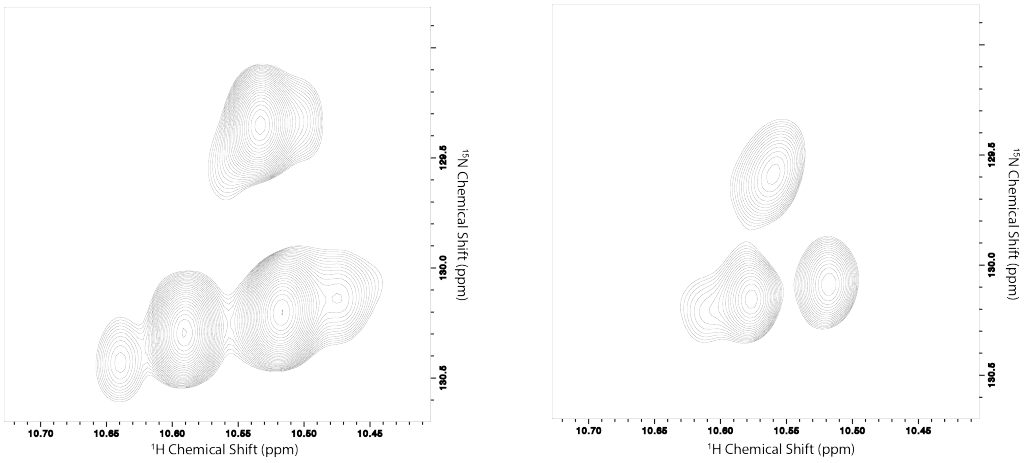


(a) DPC



(b) DPC/DPPC, $q = 0.1$

Figure 6.6: ^{1}H–^{15}N HSQC spectra of [U–^{15}N]-PDGFβR-TM-20 with and without the addition of DPPC. All experiments are conducted on 0.1 mM of protein at 37 °C in a 50 mM sodium phosphate buffer with 20 mM NaCl and pH 6.0. DPC was added to a concentration of 100 mM while DPPC was added to a concentration of 10 mM.



(a) Trp side-chain peaks in DPC

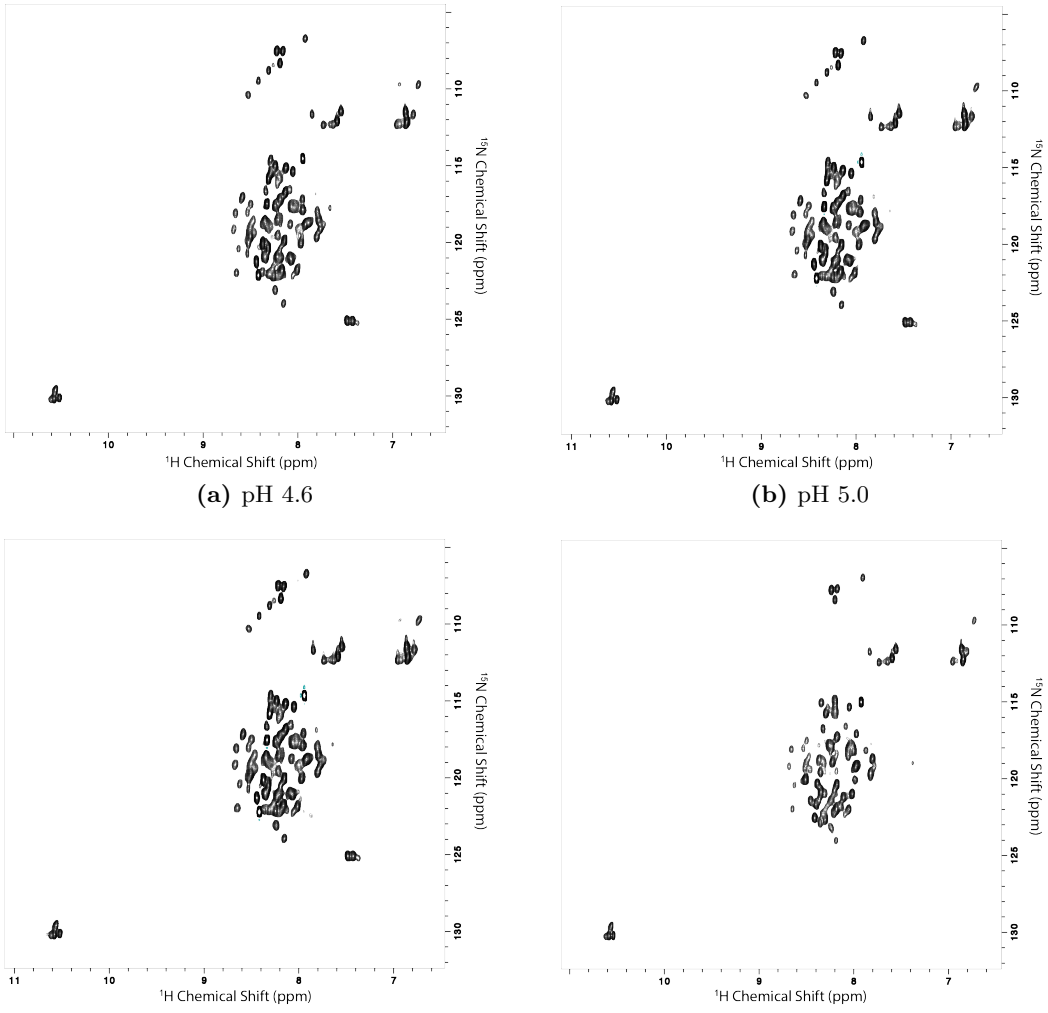
(b) Trp side-chain peaks in DPC/DPPC, $q = 0.1$

Figure 6.7: Trp side-chain peaks with and without the addition of DPPC.

6.3.6 pH

The rate of exchange between amide protons and solvent protons is pH-dependent. Although certain pulse sequences have been developed to minimise the deleterious effects of fast amide protein exchange with water (Zerbe, 2003), minimisation of the exchange rate can significantly increase sensitivity.

Four different values of pH between 4.6 and 6.8 were investigated and resulting spectra are presented in Figure 6.8. pH 4.6 yielded the best spectrum, however the protein demonstrated reduced stability at this pH. This has been observed for micellar systems previously for pH values below 6.0 (Chadwick et al., 2015) and is likely due to the pH being close to the isoelectric point (pI) at which the net charge on the protein is close to zero and the attractive forces between the proteins dominate leading to aggregation. No significant differences were observed in the spectra collected at pH 5.0 and pH 6.0. However, when increasing the pH to 6.8, several peaks disappeared from the spectrum; for example, 3 peaks at the top of the spectrum in a region which is characteristic to Gly peaks, can no longer be observed. In light of these findings, a pH of 6.0 was chosen for further studies.



(a) pH 4.6 **(b)** pH 5.0

(c) pH 6.0 **(d)** pH 6.8

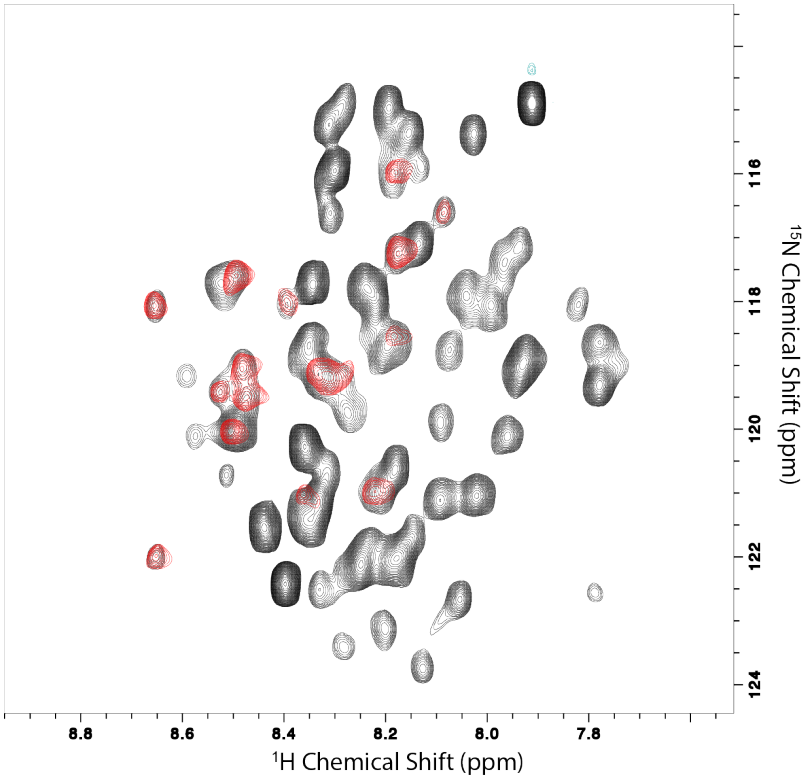
Figure 6.8: ^{1}H–^{15}N HSQC spectra of [U–^{15}N]-PDGFβR-TM-20 in DPC/DPPC at various pHs. All experiments are conducted on 0.1 mM of protein at 45 °C in a 50 mM sodium phosphate buffer with 20 mM NaCl.

6.4 Hydrogen–Deuterium Exchange Experiments

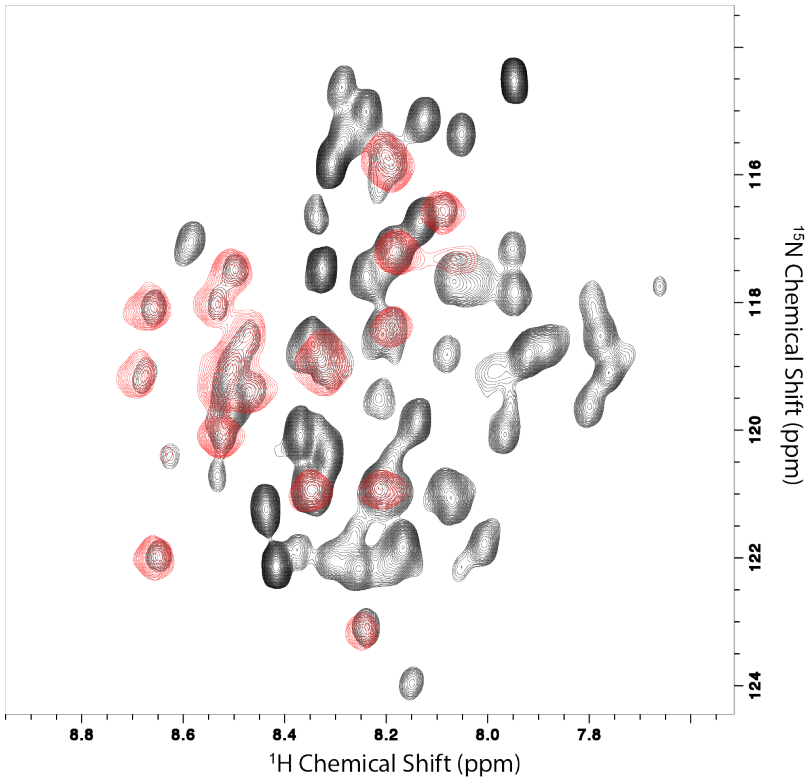
To investigate which resonances observed in the HSQC spectrum belonged to the TM region of the PDGFβR-TM-20 protein, the NMR sample was lyophilised and resuspended in 100% D_2O. This was performed for NMR samples in presence and absence of DPPC, under optimised conditions. Peaks arising from solvent-exposed protons will disappear as the protons are exchanged with deuterium. However, amide protons from transmembrane residues embedded in the (mixed) micelle are protected and will take considerably longer to exchange.

It can be seen in Figures 6.9a that only 15 peaks remain after H/D exchange in a sample prepared in DPC micelles. If DPPC is added at a $q = 0.1$, then 19 distinct peaks are observed after H/D exchange (Figure 6.9b). This is most likely because DPPC preferentially solvates the region around the transmembrane domain of the protein and it has two 16-carbon acyl chains (Figure 6.5b), compared to the single shorter 12-carbon acyl chain of DPC (Figure 6.5a).

The above would allow more residues to be embedded within the hydrophobic mixed micellar environment.



(a) DPC



(b) DPC/DPPC, $q = 0.1$

Figure 6.9: ^{1}H–^{15}N HSQC spectra of [U–^{15}N]-PDGFβR-TM-20 in (a) DPC or (b) DPC/DPPC before (black) and after (red) lyophilising the sample and resuspending in 100% D_2O. All experiments are conducted on 0.1 mM of protein at 45 °C in a 50 mM sodium phosphate buffer with 20 mM NaCl at pH 6.0.

6.5 E5 Titrations

As has been detailed in section 4.1, strong evidence exists to suggest that the bovine papillomavirus E5 protein binds to the PDGFβR-TM to form a complex. However, no studies exist to corroborate which residues are implicated in stabilising this complex. The aim here was to identify which amino acids in PDGFβR-TM play a critical role in E5 binding and to determine the effect this has on the juxtamembrane regions.

Once sample conditions had been optimised as detailed in the previous section, E5 was titrated into the PDGFβR-TM mixed micelle solution in a step-wise manner. Chemical shifts are very sensitive to the electronic environment and hence even small changes in ^{1}H and ^{15}N shifts upon the titration of a ligand can be used to map a binding interface. It was anticipated that a varying degree of chemical shift perturbation would be observed which would establish which amino acids lie close to the binding interface (e.g. T_{545} and K_{531}) and exhibit clear signs of a change in chemical environment.

The chemical shift perturbations observed upon titrating E5 to PDGFβR-TM-20 can be seen in Figure 6.10. After obtaining a ^{1}H–^{15}N HSQC spectrum of PDGFβR-TM-20, E5 was gradually added at PDGFβR-TM-20 to E5 molar ratios of $1:1$, $1:3, 1:5, 1:7, 1:9$; no further changes in chemical shifts were observed once this ratio reached $1:11$, indicative of saturation.

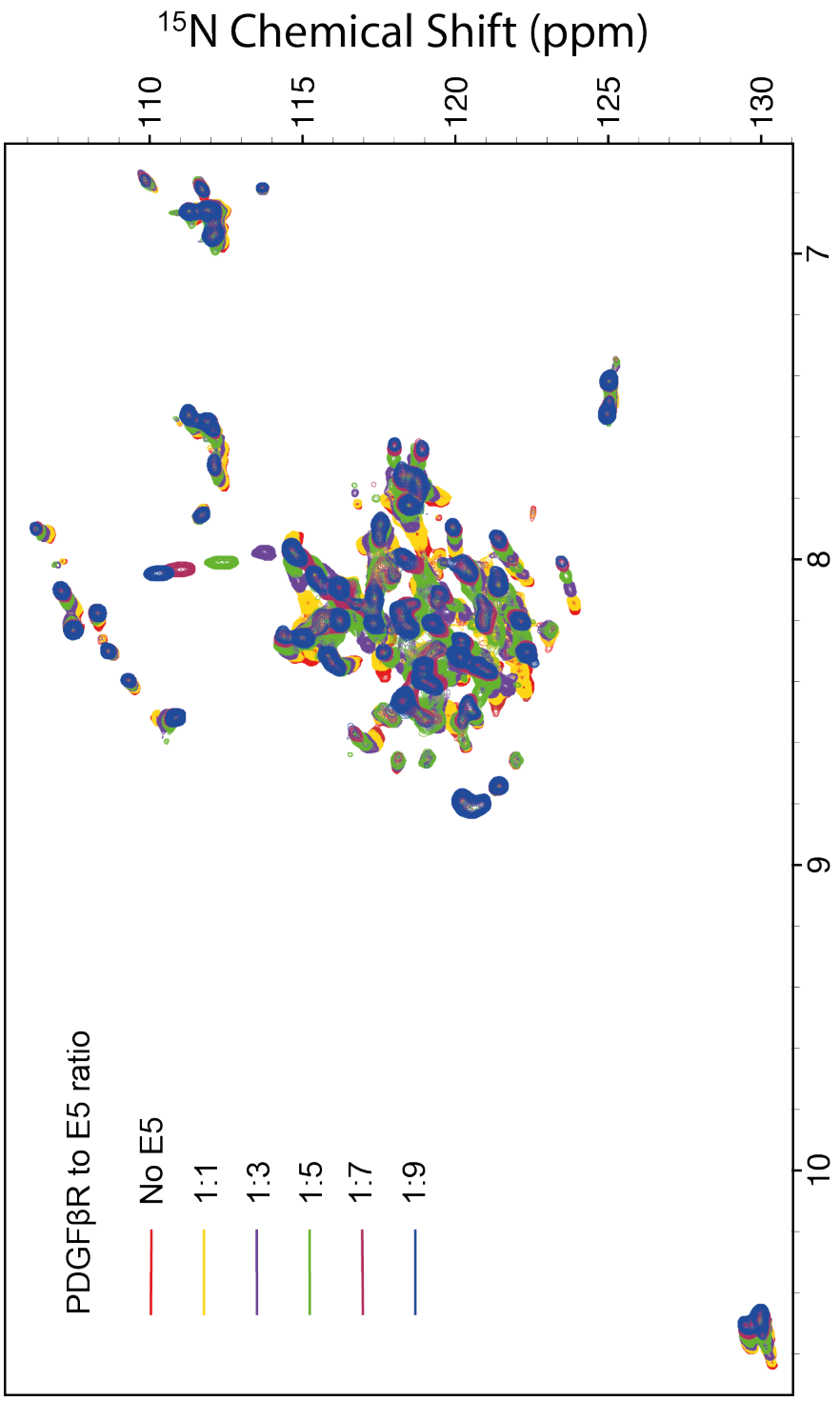


Figure 6.10: ^{1}H–^{15}N HSQC spectra of U–^{15}N-PDGFβR-TM-20 in DPC/DPPC upon the successive addition of E5.

Upon the addition of E5, a variety of phenomena were observed. There were peaks which did not exhibit any significant shift (Figure 6.11a), while a widely varying degree of shifts were observed for others (Figure 6.11b & c).

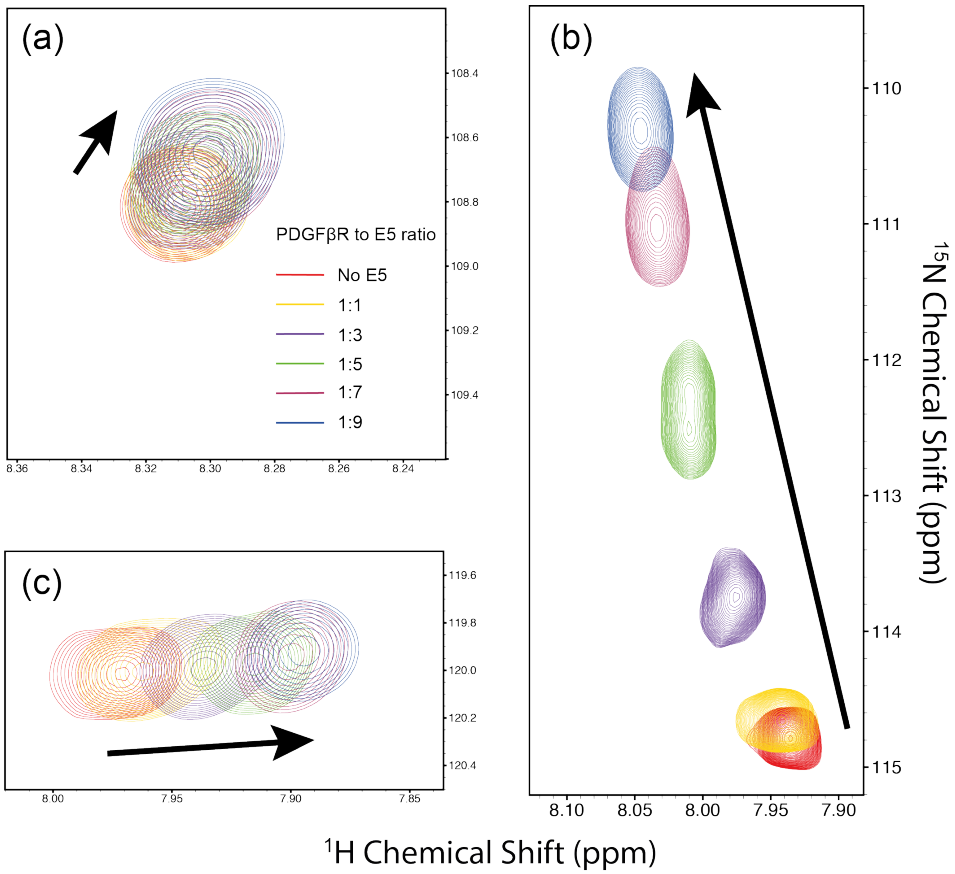
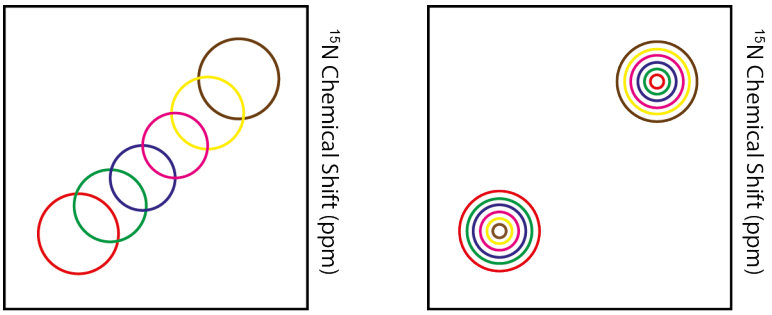


Figure 6.11: Examples of chemical shift perturbations upon the additions of E5. Some peaks did not show any significant shift (a), while a widely varying degree of shifts was observed for many others (b) and (c).

There were also peaks which disappeared upon the addition of E5, as well as the appearance of new peaks. This can be explained if the residue is involved in very strong binding to the ligand, leading to a slow rate of exchange. It has been explained previously (Williamson, 2013) that if the protein–ligand interaction is in the fast exchange regime then the chemical shifts change continuously as the binding partner is added to the protein. Conversely, the interaction could be in a slow exchange regime if peaks reduce in intensity and eventually disappear while new peaks may appear and grow in intensity simultaneously. In such a situation, the binding is so strong that it can safely be assumed that throughout the duration of the ex-

periment any one protein molecule is either free or in complex and chemical shifts for both are observed simultaneously. This has been illustrated diagrammatically in Figures 6.12a & 6.12b.



(a) Fast exchange **(b)** Slow exchange

Figure 6.12: The effect of exchange rate on NMR peak shapes. (a) Fast exchange: a gradual transition from unbound (red) to bound (brown) is seen. Slow exchange: the peak representing the unbound state (red) reduces in intensity as the peak from bound state (brown) increases in intensity. Figure adapted from (Williamson, 2013).

As an example, this phenomenon can be seen in a region of the spectrum highlighted in Figure 6.13. There are 3 peaks (encompassed in a blue box) which completely disappear upon the addition of E5 without any change in their chemical shift. Concomitantly, however, at least 3 new peaks emerge in an adjacent area of the spectrum (encompassed in a green box). Even though it could not immediately be established that the new peaks belong to the residues which have disappeared in the blue box, this observation does lend strong credence to the hypothesis that strong binding events are occurring.

Interestingly, according to the H/D exchange data presented in section 6.4, each of the 3 peaks which disappear upon the addition E5 is suggested to lie within the mixed micelle (and would thus lie in the transmembrane domain). Upon further inspection, it can be noted that 12 of the 19 peaks which putatively lie within in the mixed micelle unambiguously exhibit this property where they gradually disappear from the spectrum without much change in chemical shift. Furthermore, for all peaks that are thought to lie outside of the mixed micelle and that do not exhibit any significant change in chemical shift, their intensity is reduced only by a small

factor (if at all). This can be attributed to the the increased size of the mixed micelle.

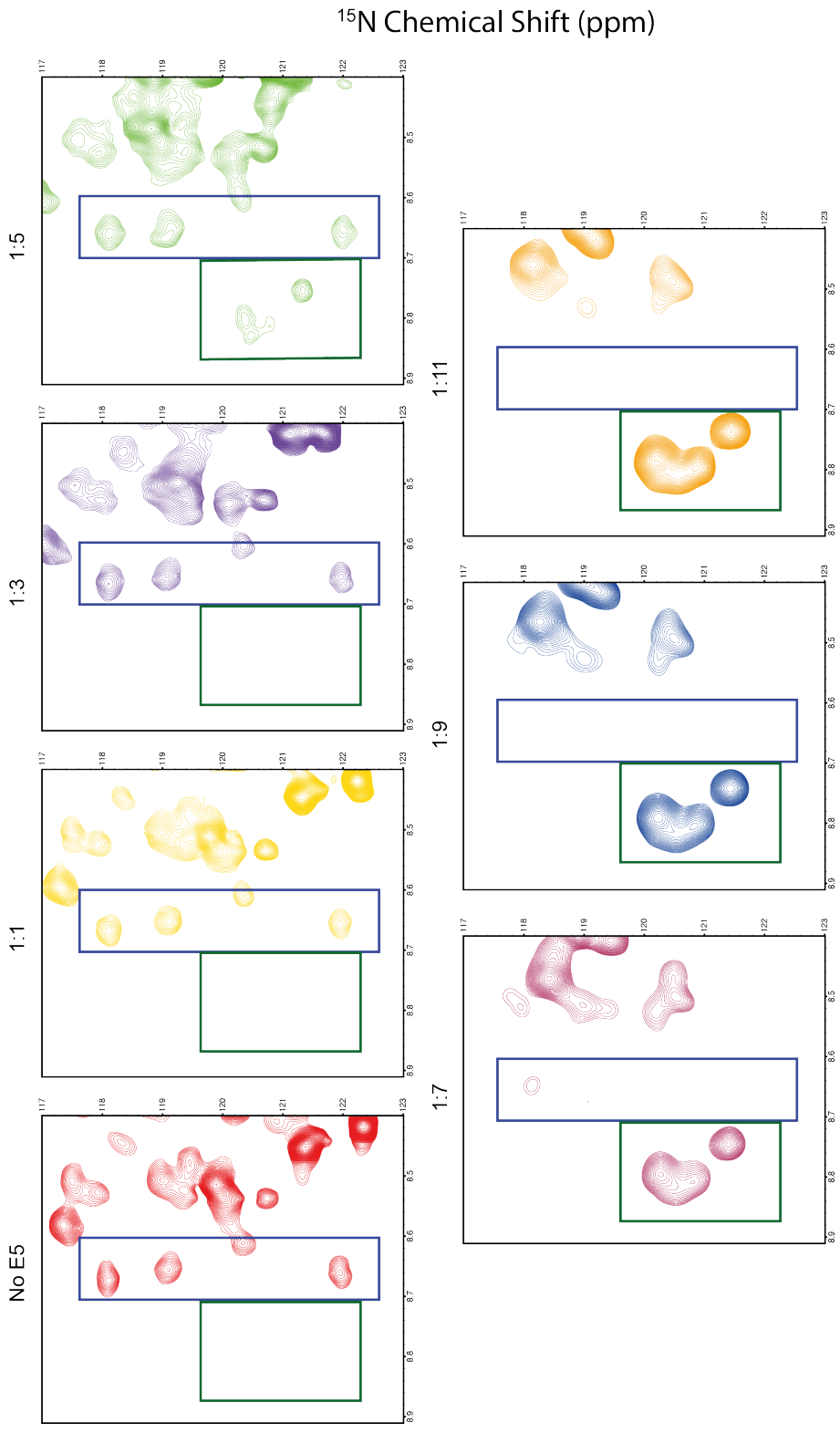


Figure 6.13: The appearance and disappearance of peaks upon the successive addition of E5.

6.6 Backbone Assignment Using triple resonance (CHN) 3D NMR

In order to extract amino acid specific information from the NMR titrations described above, it was necessary to attempt to assign each of the peaks in the ^{1}H–^{15}N HSQC spectrum. All of the triple resonance 3D NMR experiments conducted for this purpose have been introduced in section 1.5.7 and the experimental set up is detailed in section 2.14.3. The optimisations described in section 6.3 are particularly important because higher dimensional experiments have lower sensitivity, so a peak with low signal to noise may not show up at all in a 3D experiment.

Additionally, some 3D experiments are even less sensitive than others. HNCO is the most sensitive triple resonance experiment, and the relative sensitivities of other standard experiments used in this study are given in Table 6.2. The least sensitive experiments, such as CBCANH, may not allow the observation of weaker peaks in the HSQC spectrum. On the other hand, although HNCO is the most sensitive 3D experiment, the dispersion of CO nuclei is very small, ∼174–177 ppm.

Experiment	Nuclei observed	Relative S/N (%)
HNCO	H_i N_i CO_{i-1}	100
HNCA	H_i N_i $C\alpha_i$ $C\alpha_{i-1}$	50/15
HNCOCA	H_i N_i $C\alpha_{i-1}$	71
HNCACO	H_i N_i CO_i CO_{i-1}	13/4
CBCACONH	H_i N_i $C\alpha_{i-1}$ $C\beta_{i-1}$	13/9 α/β
CBCANH, HNCACB	H_i N_i $C\alpha_i$ $C\beta_i$ $C\alpha_{i-1}$ $C\beta_{i-1}$	4/1.7 α/β (i) 1.3/0.5 α/β $(i-1)$

Table 6.2: Relative sensitivity of backbone assignment experiments compared to the HNCO experiment. Adapted from (Sattler et al., 1999).

Where the complementary pairs of 3D heteronuclear experiments HNCO/HNCACO and HNCA/HNCOCA are useful in determination of backbone connectivities in amino acids of a peptide chain, the CBCACONH/CBCANH pair additionally provide useful information about amino acid type. There are 4 amino acids which give very distinct shifts and hence are easily identifiable. Gly is unique in that it is the only amino acid which gives a Cα but no Cβ signal. Furthermore, it is the only

amino acid which has an average Cα chemical shift of below 50 ppm. Ala is also uniquely placed in the spectra, with a Cβ peak which is usually below 20 ppm. Ser/Thr are the only amino acids which always show a Cβ peak at greater than 55 ppm, with Thr Cβ usually higher than 65 ppm.

6.6.1 An example of sequential backbone assignment

In order to demonstrate how sequential assignment was performed after the collection of appropriate 3D data, the N-terminal stretch of 5 amino acids VSSDG is shown in Figure 6.14. Due to the higher signal to noise, the HNCO/HNCACO and HNCA/HNCOCA datasets were simultaneously used to determine backbone connectivities as shown Figure 6.14a & b respectively. In the case of the HNCO/HNCACO pair, the HNCACO shows two peaks: one of which belongs to the CO_i (current residue) and the other CO_{i-1} (preceding residue). The HNCO, on the other hand, only shows the CO_{i-1} peaks, which allowed us to differentiate the peaks in HNCACO. Once CO_i and CO_{i-1} have been identified, the preceding residue can be located in the spectrum by looking for a residue that has a CO_i correlation at the chemical shift of CO_{i-1} correlation of this residue. This procedure is then iteratively continued. An analogous procedure is followed for the HNCA/HNCOCA pair.

While the HNCO/HNCACO pair helps trace connectivities through CO nuclei and HNCO/HNCACO does so through Cα nuclei, the CBCACONH/CBCANH can provide additional confirmation of these connectivities by tracing Cβ nuclei. As explained earlier, the Cβ's can also be used to deduce amino acid type. In the example shown in Figure 6.14, Gly can easily be identified using its distinctive Cα shift. Similarly, it can be identified that there are two consecutive Ser or Thr residues. The link X–S/T–S/T–X–G is then compared to the sequence and only one unique match was found, so the backbone assignment VSSDG was made for this group of peaks.

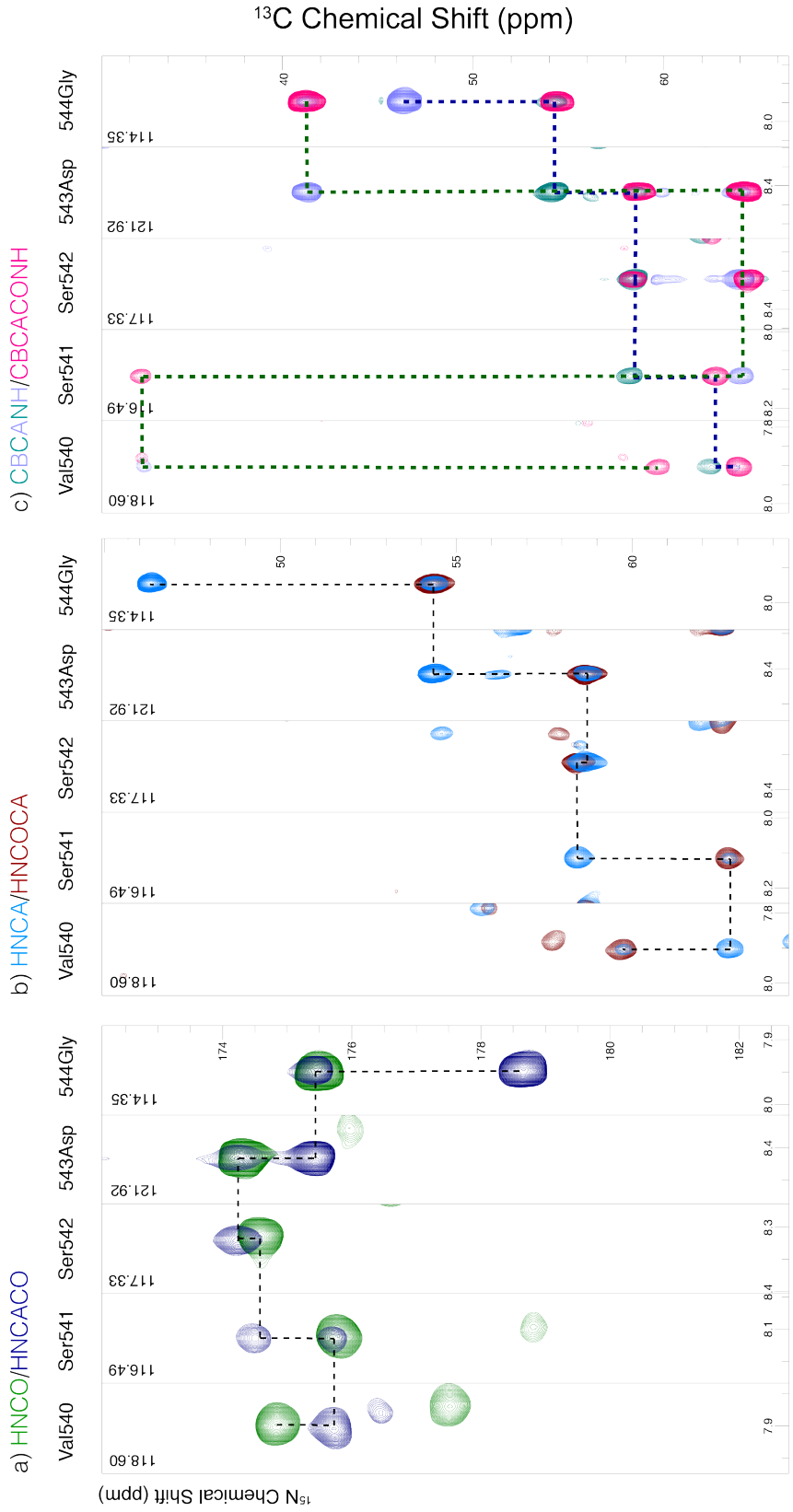


Figure 6.14: An example of sequential backbone assignment. The spectra in (a) and (b) are used for sequential assignment whereas (c) additionally provides valuable information on amino acid type.

6.6.2 Backbone assignment results

Using the above strategy, a total of 29 out of the 90 non-proline residues were assigned. The low signal to noise in other regions of the spectra inhibited further assignment. The results are shown in Figure 6.15. Any residues which were in the tag region attached to the protein are highlighted in blue. Four of the residues, 481Ser–482Met–483Gly–484Gly could be assigned to two sets of peaks, likely because they are in two different conformational states, possibly induced by the flexible glycines. Each of these is suffixed with an A or a B, where B forms a part of a continuous 14 amino acid assigned stretch. Unfortunately, no residues within the transmembrane domain could be assigned. However, 7 residues of the C-terminal juxtamembrane region and 16 residues of the N-terminal juxtamembrane region were assigned. The sequence is given below:

N-*HHHHHHGKPIPNPLLGLDSTENLYFQG*IDPLWNSMGGDSQEVTVVPHSLPFKVVVISAILALVVLTVIS
LIILIMLWQKKPRYEIRWKVIESVSSDG-C

In the sequence above, the italicised region represents the tag, and all assigned residues are underlined.

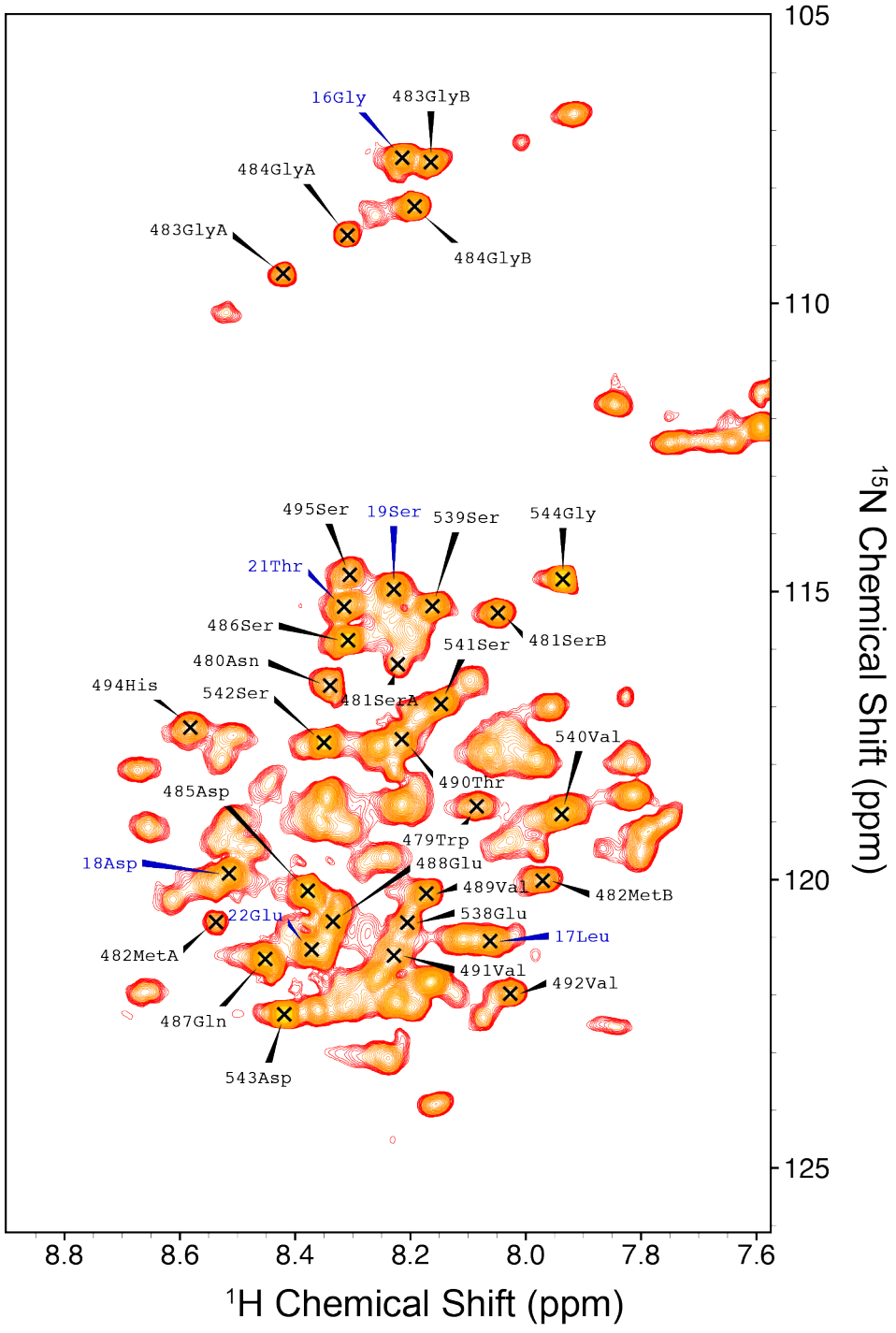
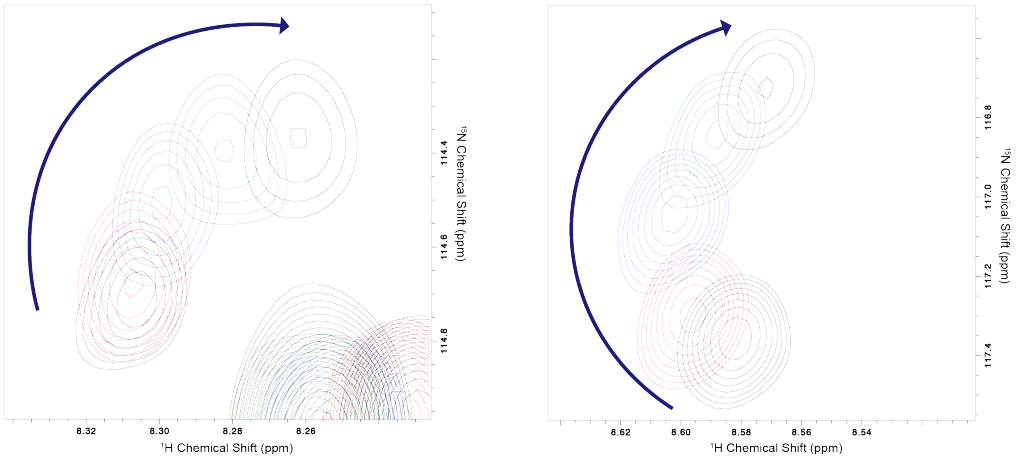


Figure 6.15: Backbone assignment of 29 residues. Assignments in blue are of residues which belong to the tag region of the expressed PDGFβR-TM-20 construct.

6.6.3 Effect of E5 on the juxtamembrane residues

Since greater than 50% of the juxtamembrane residues werre assigned, some useful information could be extrapolated. Firstly, many of the chemical shift perturbations such as those for 495Ser and 494His (Figure 6.16a & 6.16b respectively) are nonlinear which is indicative of a conformational change (Williamson, 2013). This is in line with some of the existing models of ligand binding to RTKs which predict a large conformational change in the juxtamembrane regions upon ligand binding.



(a) 495Ser **(b)** 494His

Figure 6.16: Non-linear chemical shift perturbations of (a) 495Ser and (b) 494His.

Unlike the peaks from residues proposed to lie within the micelle, the juxtamembrane peaks exhibit fast exchange and their perturbations can be followed in the NMR spectrum. In order to investigate the effects of the PDGFβR-TM/E5 interactions on the juxtamembrane residues which have been assigned, the following formula was used to calculate chemical shift perturbations (Williamson, 2013):

$$\Delta = \sqrt{\tfrac{1}{2}[\delta_H^2 + 0.14 \cdot \delta_N^2]}$$

where δ_H and δ_N are the chemical shift differences in ^{1}H and ^{15}N dimensions respectively. The results from the calculations have been illustrated in the bar graph shown on Figure 6.17.

Even though the largest shifts upon the addition of a ligand tend to be local (Williamson, 2013), when E5 is added to PDGFβR-TM-20, some very large per-

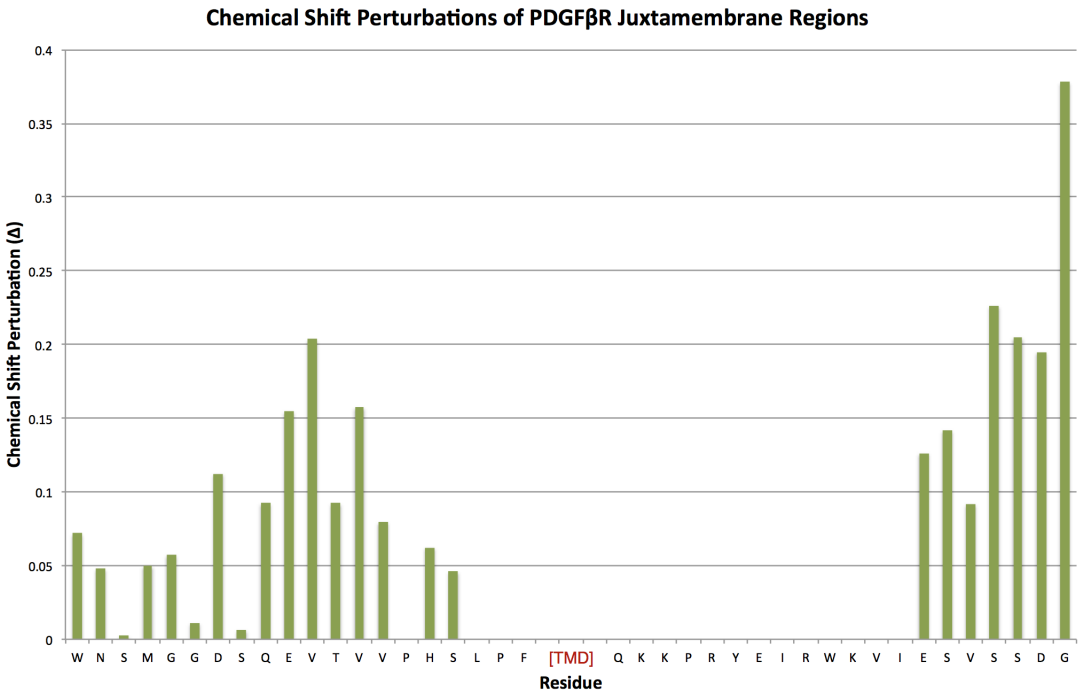


Figure 6.17: Chemical shift perturbations of PDGFβR juxtamembrane regions.

turbations are seen in the juxtamembrane domains. This is particularly evident in the last 7 residues in the C-terminus. This, however, is not unexpected as studies have shown that receptor tyrosine kinase juxtamembrane domains undergo dramatic rearrangement upon activation (Matsushita et al., 2013). Unfortunately, not enough peaks could be assigned to quantitatively assess the significance of these perturbations.

6.7 Summary and Conclusion

This chapter explained all the experiments that were carried out to investigate PDGFβR-TM/E5 interactions and the conclusions that were drawn from them. Firstly, the HSQC experiment was used to make the best possible choice of detergent and pulse sequence, and to optimise experimental conditions such as temperature and pH.

The addition of DPPC lipids yielded significant improvements to the spectrum. The number of peaks increased and narrower linewidths were also obtained; but most

importantly, it allowed higher concentrations to be reached without significant shifts in the spectrum, indicative of greater conformational stability. Hydrogen–deuterium exchange experiments also revealed that after adding DPPC, more peaks which can be attributed to the TM domain can be seen.

After the sample conditions had been optimised, E5 peptide was titrated into the sample to investigate its effects. A large variation in how much the chemical shifts were perturbed: some peaks remained at their given chemical shifts, whereas some moved significantly. It was possible follow the transitions when the rate of exchange was fast. There were several peaks which appeared to be exhibiting a slow rate of exchange given that they disappeared upon the addition of E5 without showing any chemical shift perturbations. Several new peaks were also observed, indicative of strong binding. It was additionally observed that all of the peaks which exhibited this behaviour were those which H/D experiments suggested lie within the mixed micelle.

Backbone assignment was attempted using 3D NMR experiments. It was possible to assign 29 out of the 90 non-proline residues. Even though no transmembrane domain peaks could be assigned, it was possible to do so for over half of the juxtamembrane peaks. Large shifts were observed on both sides of the transmembrane domain which was in line with receptor tyrosine kinase activation model which suggests a dramatic rearrangement of the juxtamembrane region upon activation of the receptor (Matsushita et al., 2013).

Chapter 7

Discussion and Future Work

Probing the structure of membrane proteins and the dynamics associated with various biological processes taking place within the bilayer environment is a major challenge in structural biology. This is attested by the fact that less than 3% of all entries in the Protein Data Bank (PDB) fall under the classification of membrane proteins. However, due their great potential as pharmaceutical targets, an increasing amount of resources are being expended to overcome the difficulties in the biophysical analyses of membrane proteins. This has borne some fruits in that the number of membrane protein structures determined increases several fold each year. Nonetheless, they are still disproportionately low compared to water soluble proteins.

There are several difficulties that are associated with membrane protein production and analysis due to their high hydrophobicity and requirement for reconstitution into membrane mimetics (section 1.3) if they are to be studied *in vitro*. The first major challenge is to express membrane proteins in sufficient quantities. Several of the commonly used prokaryotic expression organisms such as *E. coli* do not possess the cellular machinery to cope with post-translational processes associated with membrane protein folding during heterologous expression. Great strides have been made towards overcoming these problems which include: (1) the development of strains which are more tolerant to toxic membrane proteins like C41, C43 (Miroux and Walker, 1996b) or Lemo21 (Wagner et al., 2008); (2) tags to improve expression levels such as GFP (Drew et al., 2006) and Mistic (Roosild et al., 2005); (3) hosts like *Lactococcus lactis* for which inclusion body formation has so far not been observed (Kunji et al., 2003); (4) optimisations of eukaryotic expression systems such as insect

(Trometer and Falson, 2010) and yeast (Sreekrishna et al., 1997) cell strains; (5) preparative scale expression of membrane proteins in cell-free systems (Jermutus et al., 1998; Schwarz et al., 2007).

Once protocols have been developed to produce membrane proteins in adequate quantities, there is still the need to reconstitute them into membrane mimetics while minimising the compromise on structural integrity. Several such systems have been developed (see section 1.3) to address this challenge and the reasons which govern the specific choice of mimetic are manifold, the most crucial being its compatibility with the analytical technique employed. Sample optimisation (e.g. detergent screens) usually needs to be performed which can be laboriously time-consuming. However, techniques such as NMR are now being used to access structural information for proteins *in vivo* which represents an exciting development in a field with great potential (Gold et al., 2014; Renault et al., 2012).

The aim of this study was to establish methods for the production of, and/or use cutting edge techniques to retrieve structural information for, the transmembrane (TM) domains of receptor tyrosine kinases (RTKs) Neu and PDGFβR. As evidenced through biophysical studies conducted on the RTK-TMs, these regions play an essential role in receptor signalling and do not merely tether the receptor to the membrane. They do so in two different ways; firstly, they confer a thermodynamic stability to the TM dimers which is essential for signalling, secondly, depending on how the dimers are structurally oriented, they can maintain an active or inactive state.

Despite the medical significance of RTKs as highlighted in section 1.2, no high-resolution structures of full length RTKs are available. However, several structural studies have been performed on isolated non-TM domains of the receptors and-high resolution crystal structures of the intra- and extra-cellular regions, with and without ligands have been obtained. To the contrary, there are only two high resolution structures for RTK-TM dimers, namely, ErbB2 (Bocharov et al., 2008c) and EphA1 (Bocharov et al., 2008a). However, the role of the respective RTK-TMs in receptor signalling with respect to the conformational state in the structures is still unclear.

Furthermore, a synergistic structural understanding of how extracellular events such as ligand binding induce catalytic activity in the cytoplasm is still lacking.

Both of the aforementioned structures were obtained through solution-state NMR and there have been no successful attempts at crystallising RTK-TM dimers for structure determination. While it is hoped that high-resolution structural information of these systems will emerge following the ongoing developments in X-ray crystallography (Garman, 2014) and Cryo-electron microscopy (Goldie et al., 2014), NMR spectroscopy is uniquely placed to address the limitations of these techniques. The ability to study and understand conformational change of RTK-TM in activation of the receptor is central to developing a model of how signalling occurs. Using solution- and solid-state NMR, such conformational dynamics can be closely monitored.

Bearing the above in mind, the structure and dimeric interface of oncogenic Neu (Neu*) RTK-TM was investigated using solid-state NMR. This afforded an opportunity to carry out analyses in 'native-like' bilayers. Work conducted was based on two different models for the stabilising interactions in the Neu* homodimer. The first hypothesised $A_{661}XXXG_{665}$ motif stabilises the dimer (Smith et al., 2002b) whereas the second implicated $I_{659}XXXV_{663}$ as the stabilising interface (Beevers and Kukol, 2006; Beevers et al., 2010).

Two different samples were prepared to investigate this. For each sample, a pair of peptides, each of which was labelled at a single amino acid in the sequence, were reconstituted into DMPC/cholesterol liposomes. Solid-state NMR DARR experiments were performed to look for inter-helical correlations. The spectra obtained provided evidence to support $A_{661}XXXG_{665}$ motif as the stabilising interface due to inter-helical correlations obtained even at short mixing times. No such correlations could be observed for the other model. However, two additional conformations were observed when studying the $A_{661}XXXG_{665}$ motif; cross-correlation between two of these conformational states could suggest conformational exchange between them. High degree of mobility of the peptide was also observed when looking at the proposed $I_{659}XXXV_{663}$ interface. This could indicate more than one stabilising motif,

with one being more dominant than the others. This would fit in nicely with the rotation-coupled activation model for the Neu RTK (Bell et al., 2000) where Neu exists as dimers and activation of the receptor requires rotational positioning of the dimer such that one conformation is preferred over the other.

There are several avenues through which the work performed on the Neu peptide can be taken forward. Firstly, experiments with several more mixing times could be performed as per the methods outlined in this study. The peak intensities could then be collated and build-up curves (normalised intensity vs mixing time) drawn. The build up curves for the intra-residue and inter-residue peaks can then be compared to get a quantitative estimate for the distance between specific atoms at the dimer interface. Secondly, the method of studying singly labelled peptides can be extended further. In this study, where one sample showed cross-molecular correlations even at really low mixing times, the other showed no such correlations even at high mixing times. An approach could be adopted where labels could be incorporated into adjacent residues to the current ones and observe what effect that has. However, it would need to be ensured that experiments are feasible in terms of time and money. Thirdly, incorporation of ^{19}F nuclei into the aromatic ring of phenylalanine could be used to obtain inter-helical correlations over a longer range (10–20 Å) with very little background (Hong and Schmidt-Rohr, 2013). Fourthly, the cholesterol to lipid ratio could be further modified and a complete characterisation of the effect of this ratio on the dimeric state at various temperatures could be obtained. Fifthly, a completely new approach of uniform labelling in bacteria could be employed to obtain the complete structure of the oncogenic Neu dimer.

A substantial portion of the work presented in the preceding chapters focusses on the production, purification and structural analysis of PDGFβR RTK-TM. Problems associated with the cleavage of tags in an existing protein construct led to TOPO-cloning being used to create a new construct for the heterologous expression of TEV protease cleavable PDGFβR-TM plus 20 residues on either side. A series of expression optimisations were performed to ensure maximum yield for NMR studies. Additionally, plasmid stability was closely monitored and 'high-density' expression

methods were employed to ensure maximum protein yields. Once expression conditions had been fully optimised, two detailed protocols for the purification and reconstitution of PDGFβR-TM-20 were devised, one each for solid- and solution-state NMR analysis. For solid-state NMR, purification was performed in Triton X-100 detergent and the protein was reconstituted in liposomes using BioBeads. For solution-state NMR, IMAC purification was done in a chaotropic agent such as Guanidine HCl or Urea, another purification step involving HPLC followed and the final sample was prepared by directly dissolving the protein in detergent solution.

No immediate conclusions could be drawn from the preliminary data obtained through solid-state NMR spectroscopy and further analysis was precluded due to time constraints, however, several interesting observations came to light in solution-state NMR studies of PDGFβR-TM-20. A detailed screening of various conditions was performed to get the best possible HSQC spectrum (see section 6.2). What was particularly noteworthy was that addition of DPPC lipids to the sample led to narrower linewidths, more peaks and higher concentrations were achieved without any change in chemical shifts. Hydrogen–deuterium exchange experiments revealed a set of peaks suggesting that they are associated with residues embedded within the micelles. E5 was also titrated into the sample and a variety of changes were observed. While some peaks remained largely in their original position, others moved considerably, while yet others disappeared altogether. The disappearance of some peaks and the appearance of new peaks suggested strong binding events. Interestingly, the peaks exhibiting this phenomenon were the ones which hydrogen–deuterium exchange experiments suggested would lie within the TM domain. Even though no transmembrane domain peaks could assigned, a section of the juxtamembrane regions was assigned on both sides of the TM. During E5 titrations all of these peaks exhibited significant shifts which is in line with the hypothesis that the juxtamembrane regions are substantially rearranged following activation (Matsushita et al., 2013).

Amongst the endeavours which could immediately be undertaken to take the study

of PDGFβR-TM forward is to complete the backbone assignment. This could be done by deuterating the sample and maximising the benefits of the TROSY effect in obtaining 3D spectra for assignment. Once this is done, assignment of all new peaks after the addition of E5 could also be done to show direct evidence of E5 binding to PDGFβR-TM. This can be followed by side-chain assignment culminating in the complete structure of E5/PDGFβR-TM complex. In the longer term, lipid-juxtamembrane should also be investigated. Other membrane mimetics (section 1.3) such as bicelles are also promising candidates to be considered for high-resolution structure determination. Eventually, the length the juxtamembrane region being considered could be increased to get a holistic understanding of how signal transduction occurs across the membrane. More generally, a cross-disciplinary approach could be adopted to study RTK-TMs; for example, bicelles represent a unique system in that they can be crystallised and analysed by solution- and solid-state NMR. Moreover, methods for crystallising membrane proteins could be adopted to prepare RTK samples for solid-state NMR as they would provide greater local order and have the potential of providing high-resolution structures as has already been demonstrated (Shahid et al., 2012b).

Bibliography

Abdine, A., Verhoeven, M. A., Park, K.-H., Ghazi, A., Guittet, E., Berrier, C., Van Heijenoort, C., and Warschawski, D. E. (2010). Structural study of the membrane protein mscl using cell-free expression and solid-state nmr. *Journal of Magnetic Resonance*, 204(1):155–159.

Ali, M., De Planque, M., Huynh, N., Manolios, N., and Separovic, F. (2001). Biophysical studies of a transmembrane peptide derived from the t cell antigen receptor. *Letters in Peptide Science*, 8(3-5):227–233.

Allen, T. M., Romans, A. Y., Kercret, H., and Segrest, J. P. (1980). Detergent removal during membrane reconstitution. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 601:328–342.

Andrae, J., Gallini, R., and Betsholtz, C. (2008). Role of platelet-derived growth factors in physiology and medicine. *Genes & development*, 22(10):1276–1312.

Andrew, E., Bradbury, A., and Eades, R. (1959). Removal of dipolar broadening of nuclear magnetic resonance spectra of solids by specimen rotation.

Baker, M. (2010). Making membrane proteins for structures: a trillion tiny tweaks. *Nature methods*, 7(6):429–434.

Baneyx, F. (1999). Recombinant protein expression in escherichia coli. *Current opinion in biotechnology*, 10(5):411–421.

Bax, A. and Ikura, M. (1991). An efficient 3d nmr technique for correlating the proton and15n backbone amide resonances with the α-carbon of the preceding residue in uniformly15n/13c enriched proteins. *Journal of biomolecular NMR*, 1(1):99–104.

Bayburt, T. H. and Sligar, S. G. (2010). Membrane protein assembly into nanodiscs. *FEBS letters*, 584(9):1721–1727.

Beevers, A. J., Damianoglou, A., Oates, J., Rodger, A., and Dixon, A. M. (2010). Sequence-dependent oligomerization of the neu transmembrane domain suggests inhibition of conformational switching by an oncogenic mutant. *Biochemistry*, 49(13):2811–2820.

Beevers, A. J. and Kukol, A. (2006). The transmembrane domain of the oncogenic mutant erbb-2 receptor: a structure obtained from site-specific infrared dichroism and molecular dynamics. *Journal of molecular biology*, 361(5):945–953.

Bell, C. A., Tynan, J. A., Hart, K. C., Meyer, A. N., Robertson, S. C., and Donoghue, D. J. (2000). Rotational coupling of the transmembrane and kinase domains of the neu receptor tyrosine kinase. *Molecular biology of the cell*, 11(10):3589–3599.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, 28(1):235–242.

Bernard, P. and Couturier, M. (1992). Cell killing by the f plasmid ccdb protein involves poisoning of dna-topoisomerase {II} complexes. *Journal of Molecular Biology*, 226(3):735 – 745.

Bernard, P., Kézdy, K. E., Melderen, L. V., Steyaert, J., Wyns, L., Pato, M. L., Higgins, P. N., and Couturier, M. (1993). The f plasmid ccdb protein induces efficient atp-dependent {DNA} cleavage by gyrase. *Journal of Molecular Biology*, 234(3):534 – 541.

Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J.-B., Salvi, D., Seigneurin-Berny, D., Richaud, P., et al. (2011a). Heterologous expression of membrane proteins: choosing the appropriate host. *PLoS One*, 6(12):e29191.

Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J.-B., Salvi, D., Seigneurin-Berny, D., Richaud, P., Joyard, J., Pignol, D., Sabaty, M., Desnos, T., Pebay-Peyroula, E., Darrouzet, E., Vernet, T., and

Rolland, N. (2011b). Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host. *PLoS ONE*, 6(12):e29191.

Bertani, G. (2004). Lysogeny at mid-twentieth century: P1, p2, and other experimental systems. *Journal of bacteriology*, 186(3):595–600.

Berthold, D. A., Jeisy, V. J., Sasser, T. L., Shea, J. J., Frericks, H. L., Shah, G., and Rienstra, C. M. (2011). Top ten tips for producing 13c, 15n protein in abundance. *Stable Isotopes for Structural Biomolecular NMR*, page 51.

Blume-Jensen, P. and Hunter, T. (2001). Oncogenic kinase signalling. *Nature*, 411(6835):355–365.

Bocharov, E. V., Mayzel, M. L., Volynsky, P. E., Goncharuk, M. V., Ermolyuk, Y. S., Schulga, A. A., Artemenko, E. O., Efremov, R. G., and Arseniev, A. S. (2008a). Spatial structure and ph-dependent conformational diversity of dimeric transmembrane domain of the receptor tyrosine kinase epha1. *Journal of Biological Chemistry*, 283(43):29385–29395.

Bocharov, E. V., Mineev, K. S., Volynsky, P. E., Ermolyuk, Y. S., Tkach, E. N., Sobol, A. G., Chupin, V. V., Kirpichnikov, M. P., Efremov, R. G., and Arseniev, A. S. (2008b). Spatial Structure of the Dimeric Transmembrane Domain of the Growth Factor Receptor ErbB2 Presumably Corresponding to the Receptor Active State. *Journal of Biological Chemistry*, 283(11):6950–6956.

Bocharov, E. V., Mineev, K. S., Volynsky, P. E., Ermolyuk, Y. S., Tkach, E. N., Sobol, A. G., Chupin, V. V., Kirpichnikov, M. P., Efremov, R. G., and Arseniev, A. S. (2008c). Spatial structure of the dimeric transmembrane domain of the growth factor receptor erbb2 presumably corresponding to the receptor active state. *Journal of Biological Chemistry*, 283(11):6950–6956.

Bocharov, E. V., Pustovalova, Y. E., Pavlov, K. V., Volynsky, P. E., Goncharuk, M. V., Ermolyuk, Y. S., Karpunin, D. V., Schulga, A. A., Kirpichnikov, M. P., Efremov, R. G., et al. (2007). Unique dimeric structure of bnip3 transmembrane

domain suggests membrane permeabilization as a cell death trigger. *Journal of Biological Chemistry*, 282(22):16256–16266.

Bodenhausen, G. and Ruben, D. J. (1980). Natural abundance nitrogen-15 nmr by enhanced heteronuclear spectroscopy. *Chemical Physics Letters*, 69(1):185–189.

Brady, J. P., Claridge, J. K., Smith, P. G., and Schnell, J. R. (2015). A conserved amphipathic helix is required for membrane tubule formation by yop1p. *Proceedings of the National Academy of Sciences*, 112(7):E639–E648.

Bublil, E. M. and Yarden, Y. (2007). The egf receptor family: spearheading a merger of signaling and therapeutics. *Current opinion in cell biology*, 19(2):124–134.

Bugge, K., Steinocher, H., Brooks, A. J., Lindorff-Larsen, K., and Kragelund, B. B. (2015). Exploiting hydrophobicity for efficient production of transmembrane helices for structure determination by nmr spectroscopy. *Analytical chemistry*.

Bulheller, B. M., Rodger, A., and Hirst, J. D. (2007). Circular and linear dichroism of proteins. *Physical Chemistry Chemical Physics*, 9(17):2020–2035.

Cady, S. D., Mishanina, T. V., and Hong, M. (2009). Structure of amantadine-bound m2 transmembrane peptide of influenza a in lipid bilayers from magic-angle-spinning solid-state nmr: the role of ser31 in amantadine binding. *Journal of molecular biology*, 385(4):1127–1141.

Cady, S. D., Schmidt-Rohr, K., Wang, J., Soto, C. S., DeGrado, W. F., and Hong, M. (2010). Structure of the amantadine binding site of influenza m2 proton channels in lipid bilayers. *Nature*, 463(7281):689–692.

Cai, M., Huang, Y., Sakaguchi, K., Clore, G. M., Gronenborn, A. M., and Craigie, R. (1998). An efficient and cost-effective isotope labeling protocol for proteins expressed in shape escherichia coli. *Journal of biomolecular NMR*, 11(1):97–102.

Carpino, L. A. and Han, G. Y. (1972). 9-fluorenylmethoxycarbonyl amino-protecting group. *The Journal of Organic Chemistry*, 37(22):3404–3409.

Cascio, M. and Wallace, B. (1995). Effects of local environment on the circular dichroism spectra of polypeptides. *Analytical biochemistry*, 227(1):90–100.

Chadwick, A. C., Jensen, D. R., Peterson, F. C., Volkman, B. F., and Sahoo, D. (2015). Expression, purification and reconstitution of the c-terminal transmembrane domain of scavenger receptor bi into detergent micelles for nmr analysis. *Protein expression and purification*, 107:35–42.

Chen, Y. and Wallace, B. (1997). Secondary solvent effects on the circular dichroism spectra of polypeptides in non-aqueous environments: Influence of polarisation effects on the far ultraviolet spectra of alamethicin. *Biophysical chemistry*, 65(1):65–74.

Claridge, J. K. and Schnell, J. R. (2012). Bacterial production and solution nmr studies of a viral membrane ion channel. In *Protein NMR Techniques*, pages 165–179. Springer.

Clubb, R. T., Thanabal, V., and Wagner, G. (1992). A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1 h n, 15 n, and 13 c chemical shifts in 15 n 13 c-labelled proteins. *Journal of Magnetic Resonance (1969)*, 97(1):213–217.

Crick, D. J., Wang, J. X., Graham, B., Swarbrick, J. D., Mott, H. R., and Nietlispach, D. (2015). Integral membrane protein structure determination using pseudocontact shifts. *Journal of biomolecular NMR*, 61(3-4):197–207.

Das, N., Murray, D. T., and Cross, T. A. (2013). Lipid bilayer preparations of membrane proteins for oriented and magic-angle spinning solid-state nmr samples. *Nature protocols*, 8(11):2256–2270.

Demers, J.-P., Habenstein, B., Loquet, A., Vasa, S. K., Giller, K., Becker, S., Baker, D., Lange, A., and Sgourakis, N. G. (2014). High-resolution structure of the shigella type-iii secretion needle by solid-state nmr and cryo-electron microscopy. *Nature communications*, 5.

di Guana, C., Lib, P., Riggsa, P. D., and Inouyeb, H. (1988). Vectors that facilitate

the expression and purification of foreign peptides in escherichia coli by fusion to maltose-binding protein. *Gene*, 67(1):21–30.

Ding, Y., Fujimoto, L. M., Yao, Y., and Marassi, F. M. (2015). Solid-state nmr of the yersinia pestis outer membrane protein ail in lipid bilayer nanodiscs sedimented by ultracentrifugation. *Journal of biomolecular NMR*, 61(3-4):275–286.

Dowhan, W. and Bogdanov, M. (2011). Lipid–protein interactions as determinants of membrane protein structure and function. *Biochemical Society Transactions*, 39(3):767.

Drew, D., Lerch, M., Kunji, E., Slotboom, D.-J., and de Gier, J.-W. (2006). Optimization of membrane protein overexpression and purification using gfp fusions. *Nature methods*, 3(4):303–313.

Dvir, H. and Choe, S. (2009). Bacterial expression of a eukaryotic membrane protein in fusion to various mistic orthologs. *Protein expression and purification*, 68(1):28–33.

Elter, S., Raschle, T., Arens, S., Viegas, A., Gelev, V., Etzkorn, M., and Wagner, G. (2014). The use of amphipols for nmr structural characterization of 7-tm proteins. *The Journal of membrane biology*, 247(9-10):957–964.

Faham, S. and Bowie, J. U. (2002). Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of molecular biology*, 316(1):1–6.

Falconi, M., Brunelli, M., Pesce, A., Ferrario, M., Bolognesi, M., and Desideri, A. (2003). Static and dynamic water molecules in cu, zn superoxide dismutase. *Proteins: Structure, Function, and Bioinformatics*, 51(4):607–615.

Fernández, C. and Wüthrich, K. (2003). Nmr solution structure determination of membrane proteins reconstituted in detergent micelles. *FEBS letters*, 555(1):144–150.

Frericks, H. L., Zhou, D. H., Yap, L. L., Gennis, R. B., and Rienstra, C. M. (2006).

Magic-angle spinning solid-state nmr of a 144 kda membrane protein complex: E. coli cytochrome bo3 oxidase. *Journal of biomolecular NMR*, 36(1):55–71.

Fu, R., Wang, X., Li, C., Santiago-Miranda, A. N., Pielak, G. J., and Tian, F. (2011). In situ structural characterization of a recombinant protein in native escherichia coli membranes with solid-state magic-angle-spinning nmr. *Journal of the American Chemical Society*, 133(32):12370–12373.

Fung, B., Khitrin, A., and Ermolaev, K. (2000). An improved broadband decoupling sequence for liquid crystals and solids. *Journal of Magnetic Resonance*, 142(1):97–101.

Garman, E. F. (2014). Developments in x-ray crystallographic structure determination of biological macromolecules. *Science*, 343(6175):1102–1108.

Gauglitz, G. and Moore, D. S. (2014). *Handbook of Spectroscopy, 4 Volume Set*. John Wiley & Sons.

Ghimire, H., Abu-Baker, S., Sahu, I. D., Zhou, A., Mayo, D. J., Lee, R. T., and Lorigan, G. A. (2012). Probing the helical tilt and dynamic properties of membrane-bound phospholamban in magnetically aligned bicelles using electron paramagnetic resonance spectroscopy. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1818(3):645–650.

Gold, V. A., Ieva, R., Walter, A., Pfanner, N., van der Laan, M., and Kühlbrandt, W. (2014). Visualizing active membrane protein complexes by electron cryotomography. *Nature communications*, 5.

Goldie, K. N., Abeyrathne, P., Kebbel, F., Chami, M., Ringler, P., and Stahlberg, H. (2014). Cryo-electron microscopy of membrane proteins. In *Electron Microscopy*, pages 325–341. Springer.

Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse dnas into escherichia coli methylation-restriction mutants. *Proceedings of the National Academy of Sciences*, 87(12):4645–4649.

Green, M. R. and Sambrook, J. (2012). *Molecular cloning: a laboratory manual*, volume 1. Cold Spring Harbor Laboratory Press New York.

Grzesiek, S. and Bax, A. (1992). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*, 99(1):201–207.

Hagn, F., Etzkorn, M., Raschle, T., and Wagner, G. (2013). Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *Journal of the American Chemical Society*, 135(5):1919–1925.

Hammarström, M., Hellgren, N., van den Berg, S., Berglund, H., and Härd, T. (2002). Rapid screening for improved solubility of small human proteins produced as fusion proteins in escherichia coli. *Protein Science*, 11(2):313–321.

Hanahan, D. (1983). Studies on transformation of escherichia coli with plasmids. *Journal of molecular biology*, 166(4):557–580.

Harris, R. K., Becker, E. D., De Menezes, S. M. C., Granger, P., Hoffman, R. E., Zilm, K. W., et al. (2008). Further conventions for nmr shielding and chemical shifts iupac recommendations 2008. *Solid state nuclear magnetic resonance*, 33(3):41–56.

Heldin, C.-H. and Westermark, B. (1999). Mechanism of action and in vivo role of platelet-derived growth factor. *Physiological reviews*, 79(4):1283–1316.

Helenius, A. and Simons, K. (1975). Solubilization of membranes by detergents. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 415(1):29–79.

Hiller, M., Krabben, L., Vinothkumar, K. R., Castellani, F., van Rossum, B.-J., Kühlbrandt, W., and Oschkinat, H. (2005). Solid-state magic-angle spinning nmr of outer-membrane protein g from escherichia coli. *ChemBioChem*, 6(9):1679–1684.

Hochuli, E., Döbeli, H., and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *Journal of Chromatography A*, 411:177–184.

Holbro, T., Civenni, G., and Hynes, N. E. (2003). The erbb receptors and their role in cancer progression. *Experimental cell research*, 284(1):99–110.

Holloway, P. (1973). A simple procedure for removal of triton x-100 from protein samples. *Analytical biochemistry*, 53(1):304–308.

Hong, M. and Schmidt-Rohr, K. (2013). Magic-angle-spinning nmr techniques for measuring long-range distances in biological macromolecules. *Accounts of chemical research*, 46(9):2154–2163.

Horwitz, B., Burkhardt, A., Schlegel, R., and DiMaio, D. (1988). 44-amino-acid e5 transforming protein of bovine papillomavirus requires a hydrophobic core and specific carboxyl-terminal amino acids. *Molecular and cellular biology*, 8(10):4071–4078.

Hua, Z. and Kobertz, W. R. (2013). Chemical derivatization and purification of peptide-toxins for probing ion channel complexes. In *Chemical Neurobiology*, pages 19–30. Springer.

Hwang, P. M., Bishop, R. E., and Kay, L. E. (2004). The integral membrane enzyme pagp alternates between two dynamically distinct states. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26):9618–9623.

Hynes, N. E. and Stern, D. F. (1994). The biology of erbb-2/nue/her-2 and its role in cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1198(2):165–184.

Jamshad, M., Lin, Y., Knowles, T., Parslow, R., Harris, C., Wheatley, M., Poyner, D., Bill, R., Thomas, O. T., Overduin, M., et al. (2011). Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochemical Society transactions*, 39(3):813.

Jansson, M., Li, Y.-C., Jendeberg, L., Anderson, S., Montelione, G. T., and Nilsson, B. (1996). High-level production of uniformly 15n-and 13c-enriched fusion proteins in escherichia coli. *Journal of biomolecular NMR*, 7(2):131–141.

Jermutus, L., Ryabova, L. A., and Plückthun, A. (1998). Recent advances in pro-

ducing and selecting functional proteins by using cell-free translation. *Current opinion in biotechnology*, 9(5):534–548.

Jerpseth, M., Jerpseth, B., Briester, L., and Greener, A. (1998). *Strategies*, 11(1):3–4.

Kainosho, M. (1997). Isotope labelling of macromolecules for structural determinations. *Nature structural biology*, 4:858–861.

Kang, C. and Li, Q. (2011). Solution nmr study of integral membrane proteins. *Current opinion in chemical biology*, 15(4):560–569.

Kapust, R. B., Tözsér, J., Fox, J. D., Anderson, D., Cherry, S., Copeland, T. D., and Waugh, D. S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Engineering*, 14(12):993–1000.

Katoh, E., Takegoshi, K., and Terao, T. (2004). 13c nuclear overhauser polarization-magic-angle spinning nuclear magnetic resonance spectroscopy in uniformly 13c-labeled solid proteins. *Journal of the American Chemical Society*, 126(11):3653–3657.

Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990). Three-dimensional triple-resonance nmr spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*, 89(3):496–514.

Kelly, S. M., Jess, T. J., and Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1751(2):119–139.

Kelly, S. M. and Price, N. C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Current protein and peptide science*, 1(4):349–384.

Kim, H. J., Howell, S. C., Van Horn, W. D., Jeon, Y. H., and Sanders, C. R. (2009). Recent advances in the application of solution nmr spectroscopy to multi-span

integral membrane proteins. *Progress in nuclear magnetic resonance spectroscopy*, 55(4):335–360.

Knowles, T. J., Finka, R., Smith, C., Lin, Y.-P., Dafforn, T., and Overduin, M. (2009). Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *Journal of the American Chemical Society*, 131(22):7484–7485.

Kolodziejski, W. and Klinowski, J. (2002). Kinetics of cross-polarization in solid-state nmr: a guide for chemists. *Chemical reviews*, 102(3):613–628.

Krueger-Koplin, R. D., Sorgen, P. L., Krueger-Koplin, S. T., Rivera-Torres, I. O., Cahill, S. M., Hicks, D. B., Grinius, L., Krulwich, T. A., and Girvin, M. E. (2004). An evaluation of detergents for nmr structural studies of membrane proteins. *Journal of biomolecular NMR*, 28(1):43–57.

Kunji, E. R., Slotboom, D.-J., and Poolman, B. (2003). Lactococcus lactis as host for overproduction of functional membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1610(1):97–108.

Laemmli, U. K. et al. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage t4. *nature*, 227(5259):680–685.

Lau, T.-L., Partridge, A. W., Ginsberg, M. H., and Ulmer, T. S. (2008). Structure of the integrin $\beta 3$ transmembrane segment in phospholipid bicelles and detergent micelles. *Biochemistry*, 47(13):4008–4016.

Lee, A. G. (2004). How lipids affect the activities of integral membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1666(1):62–87.

Lee, C. and Griffin, R. (1989). Two-dimensional 1h/13c heteronuclear chemical shift correlation spectroscopy of lipid bilayers. *Biophysical journal*, 55(2):355.

Lemmon, M. A. and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *Cell*, 141(7):1117–1134.

Lévy, D., Bluzat, A., Seigneuret, M., and Rigaud, J.-L. (1990). A systematic study of

liposome and proteoliposome reconstitution involving bio-bead-mediated triton x-100 removal. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1025(2):179–190.

Lewandowski, J. R., Halse, M. E., Blackledge, M., and Emsley, L. (2015). Direct observation of hierarchical protein dynamics. *Science*, 348(6234):578–581.

Lewandowski, J. R., Paëpe, G. D., Eddy, M. T., and Griffin, R. G. (2009). 15n- 15n proton assisted recoupling in magic angle spinning nmr. *Journal of the American Chemical Society*, 131(16):5769–5776.

Li, E. and Hristova, K. (2010). Receptor tyrosine kinase transmembrane domains: Function, dimer structure and dimerization energetics. *Cell adhesion & migration*, 4(2):249–254.

Li, Y., Kijac, A. Z., Sligar, S. G., and Rienstra, C. M. (2006). Structural analysis of nanoscale self-assembled discoidal lipid bilayers by solid-state nmr spectroscopy. *Biophysical journal*, 91(10):3819–3828.

Liu, M., Mao, X.-a., Ye, C., Huang, H., Nicholson, J. K., and Lindon, J. C. (1998). Improved watergate pulse sequences for solvent suppression in nmr spectroscopy. *Journal of Magnetic Resonance*, 132(1):125–129.

Loquet, A., Giller, K., Becker, S., and Lange, A. (2010). Supramolecular interactions probed by 13c- 13c solid-state nmr spectroscopy. *Journal of the American Chemical Society*, 132(43):15164–15166.

Loquet, A., Sgourakis, N. G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S., et al. (2012). Atomic model of the type iii secretion system needle. *Nature*, 486(7402):276–279.

Luo, P. and Baldwin, R. L. (1997). Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry*, 36(27):8413–8421.

Luo, W., Cady, S. D., and Hong, M. (2009). Immobilization of the influenza a m2

transmembrane peptide in virus envelope- mimetic lipid membranes: A solid-state nmr investigation. *Biochemistry*, 48(27):6361–6368.

Maina, C. V., Riggs, P. D., Grandea, A. G., Slatko, B. E., Moran, L. S., Tagliamonte, J. A., McReynolds, L. A., et al. (1988). An escherichia coli vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene*, 74(2):365–373.

Mani, R., Tang, M., Wu, X., Buffy, J., Waring, A., Sherman, M., and Hong, M. (2006). Membrane-bound dimer structure of a β-hairpin antimicrobial peptide from rotational-echo double-resonance solid-state nmr. *Biochemistry*, 45(27):8341–8349.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989a). Overcoming the overlap problem in the assignment of proton nmr spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 hartmann-hahn-multiple quantum coherence and nuclear overhauser-multiple quantum coherence spectroscopy: application to interleukin 1. beta. *Biochemistry*, 28(15):6150–6156.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989b). Three-dimensional heteronuclear nmr of nitrogen-15 labeled proteins. *Journal of the American Chemical Society*, 111(4):1515–1517.

Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. *Journal of biomolecular NMR*, 20(1):71–75.

Maruyama, I. N. (2015). Activation of transmembrane cell-surface receptors via a common mechanism? the rotation model. *BioEssays*, 37(9):959–967.

Matsushita, C., Tamagaki, H., Miyazawa, Y., Aimoto, S., Smith, S. O., and Sato, T. (2013). Transmembrane helix orientation influences membrane binding of the intracellular juxtamembrane domain in neu receptor peptides. *Proceedings of the National Academy of Sciences*, 110(5):1646–1651.

Merril, C. R., Dunau, M. L., and Goldman, D. (1981). A rapid sensitive silver stain for polypeptides in polyacrylamide gels. *Analytical biochemistry*, 110(1):201–207.

Merril, C. R. and Pratt, M. E. (1986). A silver stain for the rapid quantitative detection of proteins or nucleic acids on membranes or thin layer plates. *Analytical biochemistry*, 156(1):96–110.

Mi, L.-Z., Grey, M. J., Nishida, N., Walz, T., Lu, C., and Springer, T. A. (2008). Functional and structural stability of the epidermal growth factor receptor in detergent micelles and phospholipid nanodiscs. *Biochemistry*, 47(39):10314–10323.

Miroux, B. and Walker, J. E. (1996a). Over-production of proteins inescherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology*, 260(3):289 – 298.

Miroux, B. and Walker, J. E. (1996b). Over-production of proteins inescherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of molecular biology*, 260(3):289–298.

Mitri, Z., Constantine, T., and O'Regan, R. (2012). The her2 receptor in breast cancer: pathophysiology, clinical use, and new advances in therapy. *Chemotherapy research and practice*, 2012.

Moffatt, B. A. and Studier, F. W. (1987). T7 lysozyme inhibits transcription by t7 rna polymerase. *Cell*, 49(2):221–227.

Monteiro, N., Martins, A., Reis, R. L., and Neves, N. M. (2014). Liposomes in tissue engineering and regenerative medicine. *Journal of The Royal Society Interface*, 11(101):20140459.

Morcombe, C. R. and Zilm, K. W. (2003). Chemical shift referencing in mas solid state nmr. *Journal of Magnetic Resonance*, 162(2):479–486.

Mörs, K., Roos, C., Scholz, F., Wachtveitl, J., Dötsch, V., Bernhard, F., and Glaubitz, C. (2013). Modified lipid and protein dynamics in nanodiscs. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(4):1222–1229.

Muhle-Goll, C., Hoffmann, S., Afonin, S., Grage, S. L., Polyansky, A. A., Windisch, D., Zeitler, M., Bürck, J., and Ulrich, A. S. (2012). Hydrophobic matching controls the tilt and stability of the dimeric platelet-derived growth factor receptor (pdgfr) β transmembrane segment. *Journal of Biological Chemistry*, 287(31):26178–26186.

Needham, D. and Evans, E. (1988). Structure and mechanical properties of giant lipid (dmpc) vesicle bilayers from 20. degree. c below to 10. degree. c above the liquid crystal-crystalline phase transition at 24. degree. c. *Biochemistry*, 27(21):8261–8269.

Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using coomassie brilliant blue g-250 and r-250. *Electrophoresis*, 9(6):255–262.

Nietlispach, D. and Gautier, A. (2011). Solution nmr studies of polytopic α-helical membrane proteins. *Current opinion in structural biology*, 21(4):497–508.

Nilson, L. A. and DiMaio, D. (1993). Platelet-derived growth factor receptor can mediate tumorigenic transformation by the bovine papillomavirus e5 protein. *Molecular and cellular biology*, 13(7):4137–4145.

Nilson, L. A., Gottlieb, R. L., Polack, G. W., and DiMaio, D. (1995). Mutational analysis of the interaction between the bovine papillomavirus e5 transforming protein and the endogenous beta receptor for platelet-derived growth factor in mouse c127 cells. *Journal of virology*, 69(9):5869–5874.

Oates, J., King, G., and Dixon, A. M. (2010). Strong oligomerization behavior of PDGFβ receptor transmembrane domain and its regulation by the juxtamembrane regions. *BBA - Biomembranes*, 1798(3):605–615.

Ohashi, R. and Takegoshi, K. (2006). Asymmetric c13–c13 polarization transfer under dipolar-assisted rotational resonance in magic-angle spinning nmr. *The Journal of chemical physics*, 125(21):214503.

Ortega-Roldan, J. L., Ossa, F., Amin, N. T., and Schnell, J. R. (2015). Solution

nmr studies reveal the location of the second transmembrane domain of the human sigma-1 receptor. *FEBS letters*, 589(5):659–665.

Orwick, M. C., Judge, P. J., Procek, J., Lindholm, L., Graziadei, A., Engel, A., Gröbner, G., and Watts, A. (2012). Detergent-free formation and physicochemical characterization of nanosized lipid–polymer complexes: Lipodisq. *Angewandte Chemie*, 124(19):4731–4735.

Park, S. H., Prytulla, S., De Angelis, A. A., Brown, J. M., Kiefer, H., and Opella, S. J. (2006). High-resolution nmr spectroscopy of a gpcr in aligned bicelles. *Journal of the American Chemical Society*, 128(23):7402–7403.

Payandeh, J., Scheuer, T., Zheng, N., and Catterall, W. A. (2011). The crystal structure of a voltage-gated sodium channel. *Nature*, 475(7356):353–358.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997). Attenuated t2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to nmr structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences*, 94(23):12366–12371.

Petti, L. M., Reddy, V., Smith, S. O., and DiMaio, D. (1997). Identification of amino acids in the transmembrane and juxtamembrane domains of the platelet-derived growth factor receptor required for productive interaction with the bovine papillomavirus e5 protein. *Journal of virology*, 71(10):7318–7327.

Picard, M., Dahmane, T., Garrigos, M., Gauron, C., Giusti, F., Le Maire, M., Popot, J.-L., and Champeil, P. (2006). Protective and inhibitory effects of various types of amphipols on the ca2+-atpase from sarcoplasmic reticulum: a comparative study. *Biochemistry*, 45(6):1861–1869.

Pike, L. J. (2005). Growth factor receptors, lipid rafts and caveolae: an evolving story. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1746(3):260–273.

Pines, A., Gibby, M., and Waugh, J. (1973). Proton-enhanced nmr of dilute spins in solids. *The Journal of Chemical Physics*, 59(2):569–590.

Poget, S. F. and Girvin, M. E. (2007). Solution nmr of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1768(12):3098–3106.

Polovinkin, V., Gushchin, I., Sintsov, M., Round, E., Balandin, T., Chervakov, P., Schevchenko, V., Utrobin, P., Popov, A., Borshchevskiy, V., et al. (2014). High-resolution structure of a membrane protein transferred from amphipol to a lipidic mesophase. *The Journal of membrane biology*, 247(9-10):997–1004.

Poole, P. and Finney, J. (1983). Hydration-induced conformational and flexibility changes in lysozyme at low water content. *International Journal of Biological Macromolecules*, 5(5):308–310.

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1976). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, (258):598–9.

Postis, V., Rawson, S., Mitchell, J. K., Lee, S. C., Parslow, R. A., Dafforn, T. R., Baldwin, S. A., and Muench, S. P. (2015). The use of smalps as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1848(2):496–501.

Rabilloud, T. (1990). Mechanisms of protein silver staining in polyacrylamide gels: A 10-year synthesis. *Electrophoresis*, 11(10):785–794.

Rasmussen, S. G., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., Fischetti, R. F., et al. (2007). Crystal structure of the human &bgr; 2 adrenergic g-protein-coupled receptor. *Nature*, 450(7168):383–387.

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009). Detergent binding explains anomalous sds-page migration of membrane proteins. *Proceedings of the National Academy of Sciences*, 106(6):1760–1765.

Reeves, P. J., Hwa, J., and Khorana, H. G. (1999). Structure and function in

rhodopsin: kinetic studies of retinal binding to purified opsin mutants in defined phospholipid–detergent mixtures serve as probes of the retinal binding pocket. *Proceedings of the National Academy of Sciences*, 96(5):1927–1931.

Renault, M., Tommassen-van Boxtel, R., Bos, M. P., Post, J. A., Tommassen, J., and Baldus, M. (2012). Cellular solid-state nuclear magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences*, 109(13):4863–4868.

Rice, A. J., Alvarez, F. J., Davidson, A. L., and Pinkett, H. W. (2014). Effects of lipid environment on the conformational changes of an abc importer. *Channels*, 8(4):327–333.

Rigaud, J.-L., Levy, D., Mosser, G., and Lambert, O. (1998). Detergent removal by non-polar polystyrene beads. *European Biophysics Journal*, 27(4):305–319.

Ritchie, T., Grinkova, Y., Bayburt, T., Denisov, I., Zolnerciks, J., Atkins, W., and Sligar, S. (2009). Chapter eleven-reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in enzymology*, 464:211–231.

Robinson, A. S. (2011). *Production of membrane proteins: Strategies for expression and isolation*. John Wiley & Sons.

Roccatano, D., Colombo, G., Fioroni, M., and Mark, A. E. (2002). Mechanism by which 2, 2, 2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study. *Proceedings of the National Academy of Sciences*, 99(19):12179–12184.

Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S., Riek, R., and Choe, S. (2005). Nmr structure of mistic, a membrane-integrating protein for membrane protein expression. *Science*, 307(5713):1317–1321.

Roy, I. and Gupta, M. N. (2004). Freeze-drying of proteins: some emerging concerns. *Biotechnology and applied biochemistry*, 39(2):165–177.

Saffman, P. and Delbrück, M. (1975). Brownian motion in biological membranes. *Proceedings of the National Academy of Sciences*, 72(8):3111–3113.

Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (2000). Nmr assignment and secondary structure determination of an octameric 110 kda protein using trosy in triple resonance experiments. *Journal of the American Chemical Society*, 122(31):7543–7548.

Sanders, C. R. and Prosser, R. S. (1998). Bicelles: a model membrane system for all seasons? *Structure*, 6(10):1227–1234.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional nmr experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 34(2):93–158.

Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *Cell*, 103(2):211–225.

Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S., Sørensen, O., and Griesinger, C. (1994). A general enhancement scheme in heteronuclear multidimensional nmr employing pulsed field gradients. *Journal of biomolecular NMR*, 4(2):301–306.

Schwarz, D., Junge, F., Durst, F., Frölich, N., Schneider, B., Reckel, S., Sobhanifar, S., Dötsch, V., and Bernhard, F. (2007). Preparative scale expression of membrane proteins in escherichia coli-based continuous exchange cell-free systems. *Nature protocols*, 2(11):2945–2957.

Seddon, A. M., Curnow, P., and Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1666(1-2):105–117.

Shahid, S. A., Bardiaux, B., Franks, W. T., Krabben, L., Habeck, M., van Rossum, B.-J., and Linke, D. (2012a). Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nature Methods*, 9(12):1212–1217.

Shahid, S. A., Bardiaux, B., Franks, W. T., Krabben, L., Habeck, M., van Rossum,

B.-J., and Linke, D. (2012b). Membrane-protein structure determination by solid-state nmr spectroscopy of microcrystals. *Nature methods*, 9(12):1212–1217.

Shenkarev, Z. O., Lyukmanova, E. N., Butenko, I. O., Petrovskaya, L. E., Paramonov, A. S., Shulepko, M. A., Nekrasova, O. V., Kirpichnikov, M. P., and Arseniev, A. S. (2013). Lipid–protein nanodiscs promote in vitro folding of transmembrane domains of multi-helical and multimeric membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(2):776–784.

Shim, A. H.-R., Liu, H., Focia, P. J., Chen, X., Lin, P. C., and He, X. (2010). Structures of a platelet-derived growth factor/propeptide complex and a platelet-derived growth factor/receptor complex. *Proceedings of the National Academy of Sciences*, 107(25):11307–11312.

Shuman, S. (1994). Novel approach to molecular cloning and polynucleotide synthesis using vaccinia dna topoisomerase. *Journal of Biological Chemistry*, 269(51):32678–84.

Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., and Panda, A. K. (2015). Protein recovery from inclusion bodies of escherichia coli using mild solubilization process. *Microbial cell factories*, 14(1):41.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using escherichia coli. *Protein Science*, 18(5):936–948.

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., and McGuire, W. L. (1987). Human breast cancer: correlation of relapse and survival with amplification of the her-2/neu oncogene. *Science*, 235(4785):177–182.

Smith, S. O., Smith, C., Shekar, S., Peersen, O., Ziliox, M., and Aimoto, S. (2002a). Transmembrane interactions in the activation of the neu receptor tyrosine kinase. *Biochemistry*, 41(30):9321–9332.

Smith, S. O., Smith, C., Shekar, S., Peersen, O., Ziliox, M., and Aimoto, S. (2002b).

Transmembrane Interactions in the Activation of the Neu Receptor Tyrosine Kinase . *Biochemistry*, 41(30):9321–9332.

Spera, S. and Bax, A. (1991). Empirical correlation between protein backbone conformation and c. alpha. and c. beta. 13c nuclear magnetic resonance chemical shifts. *Journal of the American Chemical Society*, 113(14):5490–5492.

Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J.-T., Smith, P. L., Wierschke, J. D., Subramaniam, A., and Birkenberger, L. A. (1997). Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast pichia pastoris. *Gene*, 190(1):55–62.

Studier, F. and Moffatt, B. A. (1986). Use of bacteriophage {T7} {RNA} polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, 189(1):113 – 130.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification*, 41(1):207–234.

Surti, T., Klein, O., Aschheim, K., DiMaio, D., and Smith, S. O. (1998). Structural models of the bovine papillomavirus e5 protein. *Proteins: Structure, Function, and Bioinformatics*, 33(4):601–612.

Switzer, R. C., Merril, C. R., and Shifrin, S. (1979). A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Analytical biochemistry*, 98(1):231–237.

Szeverenyi, N. M., Sullivan, M. J., and Maciel, G. E. (1982). Observation of spin exchange by two-dimensional fourier transform 13 c cross polarization-magic-angle spinning. *Journal of Magnetic Resonance (1969)*, 47(3):462–475.

Takegoshi, K., Imaizumi, T., and Terao, T. (2000). One-and two-dimensional 13 c–1 h/15 n–1 h dipolar correlation experiments under fast magic-angle spinning for determining the peptide dihedral angle φ. *Solid state nuclear magnetic resonance*, 16(4):271–278.

Takegoshi, K., Nakamura, S., and Terao, T. (2001). 13c–1h dipolar-assisted ro-

tational resonance in magic-angle spinning nmr. *Chemical Physics Letters*, 344(5):631–637.

Takegoshi, K. and Terao, T. (2002). C-13 nuclear overhauser polarization nuclear magnetic resonance in rotating solids: Replacement of cross polarization in uniformly c-13 labeled molecules with methyl groups. *Journal of Chemical Physics*, 117(4):1700–1707.

Talbert-Slagle, K. and DiMaio, D. (2009). The bovine papillomavirus e5 protein and the pdgf β receptor: It takes two to tango. *Virology*, 384(2):345–351.

Tan, M. and Yu, D. (2007). Molecular mechanisms of erbb2-mediated breast cancer chemoresistance. In *Breast Cancer Chemosensitivity*, pages 119–129. Springer.

Tanford, C. and Reynolds, J. A. (1976). Characterization of membrane proteins in detergent solutions. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 457(2):133–170.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, 76(9):4350–4354.

Tribet, C., Audebert, R., and Popot, J.-L. (1996). Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proceedings of the National Academy of Sciences*, 93(26):15047–15050.

Trometer, C. and Falson, P. (2010). Mammalian membrane protein expression in baculovirus-infected insect cells. In *Heterologous Expression of Membrane Proteins*, pages 105–117. Springer.

Tropea, J. E., Cherry, S., and Waugh, D. S. (2009). Expression and purification of soluble his6-tagged tev protease. In *High Throughput Protein Expression and Purification*, pages 297–307. Springer.

Tugarinov, V., Choy, W.-Y., Orekhov, V. Y., and Kay, L. E. (2005). Solution nmr-derived global fold of a monomeric 82-kda enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, 102(3):622–627.

Tzitzilonis, C., Eichmann, C., Maslennikov, I., Choe, S., and Riek, R. (2013). Detergent/nanodisc screening for high-resolution nmr studies of an integral membrane protein containing a cytoplasmic domain. *PloS one*, 8(1):e54378.

Ullrich, A. and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. *Cell*, 61(2):203–212.

Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., et al. (2008). Biomagresbank. *Nucleic acids research*, 36(suppl 1):D402–D408.

Villinger, S., Briones, R., Giller, K., Zachariae, U., Lange, A., de Groot, B. L., Griesinger, C., Becker, S., and Zweckstetter, M. (2010). Functional dynamics in the voltage-dependent anion channel. *Proceedings of the National Academy of Sciences*, 107(52):22546–22551.

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005). The ccpn data model for nmr spectroscopy: development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics*, 59(4):687–696.

Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., et al. (2008). Tuning escherichia coli for membrane protein overexpression. *Proceedings of the National Academy of Sciences*, 105(38):14371–14376.

Wallace, B., Lees, J., Orry, A., Lobley, A., and Janes, R. W. (2003). Analyses of circular dichroism spectra of membrane proteins. *Protein Science*, 12(4):875–884.

Wang, Y. and Jardetzky, O. (2002). Probability-based protein secondary structure identification using combined nmr chemical-shift data. *Protein Science*, 11(4):852–861.

Ward, M. E., Wang, S., Munro, R., Ritz, E., Hung, I., Gorkov, P. L., Jiang, Y., Liang, H., Brown, L. S., and Ladizhansky, V. (2015). In situ structural studies

of anabaena sensory rhodopsin in the e. coli membrane. *Biophysical journal*, 108(7):1683–1696.

Warschawski, D. (2013). Membrane proteins of known structure determined by nmr.

Warschawski, D. E., Arnold, A. A., Beaugrand, M., Gravel, A., Chartrand, É., and Marcotte, I. (2011). Choosing membrane mimetics for nmr structural studies of transmembrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1808(8):1957–1974.

Whitmore, L. and Wallace, B. (2004). Dichroweb, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic acids research*, 32(suppl 2):W668–W673.

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1999). *Protein identification and analysis tools in the ExPASy server*. Springer.

Williams, L. T. (1989). Signal transduction by the platelet-derived growth factor receptor. *Science*, 243(4898):1564.

Williamson, M. P. (2013). Using chemical shift perturbation to characterise ligand binding. *Progress in nuclear magnetic resonance spectroscopy*, 73:1–16.

Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995). 1h, 13c and 15n chemical shift referencing in biomolecular nmr. *Journal of biomolecular NMR*, 6(2):135–140.

Wishart, D. S. and Sykes, B. D. (1994a). The 13c chemical-shift index: a simple method for the identification of protein secondary structure using 13c chemical-shift data. *Journal of biomolecular NMR*, 4(2):171–180.

Wishart, D. S. and Sykes, B. D. (1994b). Chemical shifts as a tool for structure determination. *Methods in enzymology*, 239:363.

Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991). Relationship between nu-

clear magnetic resonance chemical shift and protein secondary structure. *Journal of molecular biology*, 222(2):311–333.

Wittekind, M. and Mueller, L. (1993). Hncacb, a high-sensitivity 3d nmr experiment to correlate amide-proton and nitrogen resonances with the alpha-and beta-carbon resonances in proteins. *Journal of Magnetic Resonance, Series B*, 101(2):201–205.

Wu, X., Wu, D., Lu, Z., Chen, W., Hu, X., and Ding, Y. (2009). A Novel Method for High-Level Production of TEV Protease by Superfolder GFP Tag. *Journal of Biomedicine and Biotechnology*, 2009(1):1–8.

Wu, X., Wu, D., Lu, Z., Chen, W., Hu, X., and Ding, Y. (2010). A novel method for high-level production of tev protease by superfolder gfp tag. *BioMed Research International*, 2009.

Yamamoto, K., Caporini, M. A., Im, S.-C., Waskell, L., and Ramamoorthy, A. (2015). Cellular solid-state nmr investigation of a membrane protein using dynamic nuclear polarization. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1848(1):342–349.

Yang, Y., Yuzawa, S., and Schlessinger, J. (2008). Contacts between membrane proximal regions of the pdgf receptor ectodomain are required for receptor activation but not for receptor dimerization. *Proceedings of the National Academy of Sciences*, 105(22):7681–7686.

Yang, Z., Zhang, L., Zhang, Y., Zhang, T., Feng, Y., Lu, X., Lan, W., Wang, J., Wu, H., Cao, C., et al. (2011). Highly efficient production of soluble proteins from insoluble inclusion bodies by a two-step-denaturing and refolding method. *PLoS One*, 6(7):e22981.

Yıldırım, M. A., Goh, K.-I., Cusick, M. E., Barabási, A.-L., and Vidal, M. (2007). Drugtarget network. *Nature biotechnology*, 25(10):1119–1126.

Yokogawa, M., Kobashigawa, Y., Yoshida, N., Ogura, K., Harada, K., and Inagaki, F. (2012). Nmr analyses of the interaction between the fyve domain of early endo-

some antigen 1 (eea1) and phosphoinositide embedded in a lipid bilayer. *Journal of Biological Chemistry*, 287(42):34936–34945.

Zerbe, O. (2003). Bionmr in drug research (methods and principles in medicinal chemistry).

Zhang, X., Gureasko, J., Shen, K., Cole, P. A., and Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell*, 125(6):1137–1149.